

ESCUELA TÉCNICA SUPERIOR DE INGENIEROS  
INDUSTRIALES Y DE TELECOMUNICACIÓN

UNIVERSIDAD DE CANTABRIA



***Trabajo Fin de Grado***

**Aplicación de sistemas de imagen  
hiperespectral para el estudio de productos  
cárnicos**

**(Application of hyperspectral imaging systems for  
the study of meat products)**

Para acceder al Título de

***Graduado en***

***Ingeniería de Tecnologías de Telecomunicación***

Autor: Alba Martín de la Torre

Octubre 2019



Trabajo de Fin de Grado

Grupo de Ingeniería Fotónica



## **GRADUADO EN INGENIERÍA DE TECNOLOGÍAS DE TELECOMUNICACIÓN**

### **CALIFICACIÓN DEL TRABAJO FIN DE GRADO**

**Realizado por: Alba Martín de la Torre**

**Director del TFG: Olga María Conde Portilla**

**Título:** “Aplicación de sistemas de imagen hiperespectral para el estudio de productos cárnicos”

**Title:** “Application of hyperspectral imaging systems for the study of meat products “

**Presentado a examen el día: 24 de Octubre de 2019**

para acceder al Título de

### **GRADUADO EN INGENIERÍA DE TECNOLOGÍAS DE TELECOMUNICACIÓN**

Composición del Tribunal:

Presidente (Apellidos, Nombre): Fernández Ibáñez, Tomás

Secretario (Apellidos, Nombre): Conde Portilla, Olga María

Vocal (Apellidos, Nombre): Valle López, Luis

Este Tribunal ha resuelto otorgar la calificación de: .....

Fdo.: El Presidente

Fdo.: El Secretario

Fdo.: El Vocal

Fdo.: El Director del TFG  
(sólo si es distinto del Secretario)

Vº Bº del Subdirector

Trabajo Fin de Grado Nº  
(a asignar por Secretaría)



# APLICACIÓN DE SISTEMAS DE IMAGEN HIPERESPECTRAL PARA EL ESTUDIO DE PRODUCTOS CÁRNICOS

Una aplicación de diagnóstico óptico para la industria agroalimentaria enfocada en la mejora de la caracterización y análisis de calidad de productos cárnicos.



**Autor:** Alba Martín de la Torre

**Director:** Olga M<sup>a</sup> Conde Portilla



## Agradecimientos

En primer lugar, me gustaría dar las gracias al Grupo de Ingeniería Fotónica por haberme dado la oportunidad de llevar a cabo este trabajo. Especialmente a José y Arturo por la ayuda y el apoyo que me han proporcionado desde el primer momento.

A todos los que empezaron siendo simplemente compañeros de clase, con los que he vivido tantos y tantos momentos durante la carrera, y han terminado siendo grandes amigos. Muchas gracias por compartir parte del camino conmigo, hacerlo mucho más ameno y ser ahora parte de mi día a día.

A mis amigos y toda la gente que ha estado a mi lado, animándome durante este proceso. Gracias por haber creído en mí, en especial a Raquel y Samu.

A Javi, por ser mi apoyo incondicional desde siempre y a Carla, que no lo habrá sido desde hace tanto tiempo, pero tienen los dos la misma culpa de que lo haya conseguido. Gracias por toda la ayuda, el cariño y hacerme ver que sí podía.

Por último y uno de los agradecimientos más grandes a Olga, como directora de este Trabajo de Fin de Grado, por el constante apoyo, por toda la confianza depositada en mí, la paciencia y la gran ayuda recibida.



## Resumen

El objetivo de este estudio es conocer la composición de diferentes productos cárnicos centrándose fundamentalmente en poder establecer los porcentajes de grasa y carne a través de sistemas de espectroscopía de imagen HSI de espejo rotativo, gracias a las diferentes técnicas de procesamiento utilizadas para el análisis de la imagen hiperespectral.

Las técnicas de procesamiento implementadas se pueden dividir en dos grupos según el tipo de clasificación que realicen. La clasificación cuantitativa de las mezclas se lleva a cabo mediante el método clásico Least Squares con el objetivo de conseguir una aproximación lineal por mínimos cuadrados de los valores de las medidas obtenidos para cada longitud de onda, y el método PLS basado en procesos estadísticos de vectores propios.

La clasificación cualitativa se consigue en este trabajo a través de diferentes técnicas, una de ellas el algoritmo Otsu aplicado sobre varios procesos de obtención de datos representativos del espectro, como son la proyección de máxima y mínima intensidad (MIP e INVMIP respectivamente) y el promedio, permite una certificación de la clasificación cualitativa. Las otras dos formas de clasificación se basan en el cálculo de distancias euclídeas y ángulos espectrales dando como resultado un mapa de la distribución espacial de los componentes.

A partir de los datos proporcionados por las diferentes técnicas de procesamiento se compararán los resultados obtenidos entre los diferentes tipos de carne y los dos sistemas HSI que trabajan en diferentes rangos del espectro, en el visible y parte del infrarrojo cercano (VISNIR) y el otro sistema en el infrarrojo cercano (NIR).



## Abstract

This study aims to determine the composition of different meat products, focusing on establish fat and meat percentages using rotating mirror HSI image spectroscopy and different processing techniques for hyperspectral image.

Implemented processing techniques can be divided into two main groups depending on the type of classification. Quantitative classification of the mixtures is done using classical Least Squares method, aiming to achieve a lineal approximation by least squares of the obtained measured values for each wavelength, and PLS approach using eigenvectors statistics.

Qualitative classification can be achieved through different techniques. Otsu algorithm is one of them, applied to various processes for obtaining representative spectrum data, such as the maximum and minimum intensity projection (MIP and INVMIP respectively) and the average, allows a qualitative classification certification. The other two ways of classification are based on the calculation of Euclidean distances and spectral angles, resulting in a map of the spatial distribution of the components.

From the data provided by the different processing techniques, the results obtained between the different types of meat and the two HSI systems that work in different ranges of the spectrum, in the visible and near infrared part (VISNIR) and the other near infrared system (NIR), will be compared.

## Palabras clave

- Imagen Hiperespectral
- VIS-NIR
- SWIR (Short Wave Infrared)
- Coeficiente de correlación
- Picos de absorción
- Reflectancia
- Método de aproximación lineal de Mínimos cuadrados, Classical Least Squares (CLS)
- Modelo de Regresión de Mínimos Cuadrados (PLS)
- Algoritmo Otsu
- Maximum Intensity Projection (MIP)
- Distancia euclídea
- Mapeo del Ángulo Espectral (SAM)



# Índice

Parte I: Introducción .....	14
Capítulo 1.....	15
Introducción y motivación del trabajo.....	15
1.1 Análisis de control y calidad en la industria alimentaria .....	15
1.2 Espectroscopía de infrarrojo cercano (NIRS) en la industria alimentaria .....	19
1.3 Objetivos de este trabajo .....	21
1.4 Estructura del documento.....	22
Parte II: Teoría y estado del arte .....	24
Capítulo 2.....	25
Propiedades espectrales de la carne.....	25
2.1 Fabricación de hamburguesas .....	25
2.2 Propiedades ópticas de los tejidos .....	26
2.3 Picos de absorción de los principales cromóforos de la carne .....	29
Capítulo 3.....	34
Sistemas espectroscópicos.....	34
3.1 Sistemas HSI.....	34
Parte III: Materiales y métodos.....	39
Capítulo 4.....	40
Materiales.....	40
4.1 Sistema HSI de espejo rotativo .....	40
4.2 Preparación de las muestras.....	45
Capítulo 5.....	49
Métodos .....	49
5.1 Técnicas de procesado.....	50
5.1.1 Técnicas de procesado para clasificación cuantitativa .....	50
5.1.2 Técnicas de procesado para clasificación cualitativa.....	52
Parte IV: Resultados .....	55
Capítulo 6.....	56





Clasificación cuantitativa .....	56
6.1. Método Clásico de Mínimos Cuadrados .....	56
6.2. Método de Vectores Propios. PLS.....	72
Capítulo 7.....	82
Clasificación cualitativa.....	82
7.1 Clasificación por medio de OTSU .....	82
7.1.1 Carne de novilla.....	83
7.1.2 Carne de cerdo.....	89
7.2 Clasificación mediante distancia euclídea.....	95
7.2.1 Carne de novilla.....	95
7.2.2 Carne de cerdo.....	98
7.3 Clasificación mediante Mapeo del Ángulo Espectral ( <i>SAM</i> ) .....	100
7.3.1 Carne de novilla.....	100
7.3.2 Carne de cerdo.....	102
Parte V: Conclusiones y líneas futuras .....	104
Capítulo 8.....	105
Conclusiones generales y líneas futuras de investigación .....	105
8.1 Conclusiones generales de este trabajo.....	105
8.2 Líneas futuras de trabajo .....	106
Bibliografía .....	108

## Índice de figuras

Figura 1. Vibraciones moleculares por radiación luminosa .....	20
Figura 2: Coeficientes de absorción de los componentes de un tejido genérico en el rango del espectro de 300 a 1200 nm [22] .....	29
Figura 3: Picos de absorción de la hemoglobina [23] .....	30
Figura 4: Coeficiente de absorción normalizado del colágeno, lípidos, hemoglobina y agua [24] ..	32
Figura 5: Coeficientes de absorción de la proteína [27] .....	33
Figura 6: Coeficientes de absorción de la grasa de cerdo [28] .....	33
Figura 7: (a) Diferencias entre imagen monocromática, RGB, multiespectral e hiperspectral. Representación gráfica de espectroscopía. (b) Diferencias entre la representación espectral de la imagen hiperspectral y la intensidad de una imagen RGB. [29] .....	35
Figura 8: Funcionamiento de las técnicas de adquisición de imagen hiperspectral [32] .....	38
Figura 9: Esquema del funcionamiento de los sistemas HSI de espejo rotativo [31] .....	41
Figura 10: Esquema general del sistema. (A) Componentes principales. (B) Sistema de escáner del espejo rotativo. [31] .....	42
Figura 11: Esquema de un espectrógrafo con un componente PGP [35] .....	44
Figura 12: Cantidad de carne de novilla (izquierda) y tocino (derecha) utilizados en la elaboración de la muestra de novilla con un 75% de carne .....	45
Figura 13: Proceso de colocación de las muestras con distintos porcentajes de grasa en la cubeta .....	46
Figura 14: Variación del porcentaje de grasa en las muestras de carne de cerdo. A 100% de carne, B 95% de carne, C 90% de carne, D 85% de carne, E 80% de carne, F 75% de carne, G 50% de carne y H 25% de carne .....	46
Figura 15: Distribución de las distintas mezclas en la cubeta tras efectuar las medidas. Comparar con la Figura 13 para observar la diferencia en el estado de la carne antes y después de la realización de las medidas .....	47
Figura 16: Distribución ROIs.(Superior) Sistema VIS-NIR. (Inferior) Sistema SWIR .....	49
Figura 17: Espectros promedio de las muestras de carne de novilla con distintos porcentajes de grasa .....	57
Figura 18: Espectros promedio de las muestras de carne de cerdo con distintos porcentajes de grasa .....	57
Figura 19: Espectros promedio de las muestras de novilla 100% de carne y cerdo 100% de carne .....	58
Figura 20: Espectros promedio y coeficiente de correlación de las muestras de carne de novilla con distintos porcentajes de grasa .....	60
Figura 21: Espectros promedio y coeficiente de correlación de las muestras de carne de novilla con distintos porcentajes de grasa (logaritmo) .....	61
Figura 22: Espectros promedio y coeficiente de correlación de las muestras de carne de cerdo con distintos porcentajes de grasa .....	62



Figura 23: Espectros promedio y coeficiente de correlación de las muestras de carne de cerdo con distintos porcentajes de grasa (logaritmo) .....	63
Figura 24: Comparativa de linealidad a 580nm pico absorción HbO <sub>2</sub> en las carnes de novilla y cerdo.....	64
Figura 25: Comparativa de linealidad a 580nm pico absorción HbO <sub>2</sub> en logaritmo de las carnes de novilla y cerdo .....	65
Figura 26: Comparativa de linealidad a 1040nm pico absorción de lípidos en de las carnes de novilla con y sin muestra al 90% .....	66
Figura 27: Comparativa de linealidad a 1040nm pico absorción de lípidos en logaritmo de la carne de novilla con y sin muestra al 90%.....	67
Figura 28. Gráficas de linealidad de las longitudes de onda donde los coeficientes son máximos en la región VIS-NIR como SWIR carne de novilla .....	68
Figura 29: Comparativa de linealidad a 762nm pico absorción de la grasa en la carne de cerdo con y sin muestra al 85%.....	69
Figura 30: Comparativa de linealidad a 762nm pico absorción de la grasa en la carne de cerdo en logaritmo con y sin muestra al 85% .....	70
Figura 31 Gráficas de linealidad de las longitudes de onda donde los coeficientes son máximos en la región VIS-NIR como SWIR carne de cerdo.....	71
Figura 32: Distribución de las zonas de entrenamiento y validación en la carne.....	72
Figura 33: PLS en región VIS-NIR sobre carne de novilla .....	73
Figura 34: PLS en región VIS-NIR sobre carne de novilla (logaritmo) .....	75
Figura 35: PLS en región VIS-NIR sobre carne de cerdo .....	76
Figura 36: PLS en región SWIR sobre carne de novilla .....	77
Figura 37: PLS en región SWIR sobre carne de cerdo.....	78
Figura 38: PLS en región VIS-NIR y SWIR sobre carne de novilla .....	80
Figura 39: PLS en región VIS-NIR y SWIR sobre carne de cerdo .....	81
Figura 40: Histogramas del algoritmo Otsu sobre carne de novilla en VIS-NIR.....	83
Figura 41: Máscaras en carne y grasa de otsu en novilla VIS-NIR.....	84
Figura 42: Comparación grasa y carne novilla en VIS-NIR usando promedio, MIP e invMIP.....	85
Figura 43: Comparación grasa novilla en VIS-NIR frente al valor de referencia .....	86
Figura 44: Histogramas del algoritmo Otsu sobre carne de novilla en SWIR.....	87
Figura 45: Máscaras en carne y grasa de otsu en novilla SWIR.....	88
Figura 46: Comparación grasa novilla en SWIR frente al valor de referencia .....	89
Figura 47: Histogramas del algoritmo otsu sobre carne de cerdo en VIS-NIR.....	90
Figura 48: Máscaras en carne y grasa de otsu en cerdo VIS-NIR.....	91
Figura 49: Comparación grasa cerdo en VIS-NIR frente al valor de referencia .....	92
Figura 50: Histogramas del algoritmo otsu sobre carne de cerdo en SWIR.....	93
Figura 51: Máscaras en carne y grasa de otsu en cerdo SWIR.....	93
Figura 52: Comparación grasa cerdo en SWIR frente al valor de referencia .....	94
Figura 53: Espectros recortados en novilla VIS-NIR para distancia euclídea.....	96
Figura 54: Clasificación cualitativa en novilla VIS-NIR mediante distancia euclídea.....	97



Figura 55: Espectros recortados en novilla SWIR para distancia euclídea .....	97
Figura 56: Clasificación cualitativa en novilla SWIR mediante distancia euclídea.....	98
Figura 57: Clasificación cualitativa en cerdo VIS-NIR mediante distancia euclídea.....	99
Figura 58: Clasificación cualitativa en cerdo SWIR mediante distancia euclídea.....	100
Figura 59: Clasificación cualitativa en novilla VIS-NIR mediante SAM .....	101
Figura 60: Clasificación cualitativa en novilla SWIR mediante SAM .....	102
Figura 61: Clasificación cualitativa en cerdo VIS-NIR mediante SAM .....	102
Figura 62: Clasificación cualitativa en cerdo SWIR mediante SAM.....	103



## Índice de tablas

Tabla 1: Diferencias entre las técnicas de adquisición de imagen hiperespectral [31] .....	37
Tabla 2: Especificaciones técnicas óptica de enfoque sistema HSI rango VISNIR [33] .....	42
Tabla 3: Especificaciones técnicas cámara AVT Mako G 223B NIR [33] .....	43
Tabla 4: Especificaciones técnicas espectrógrafo sistema HSI rango VISNIR [33] .....	44
Tabla 5: Porcentajes de grasa y de los diferentes tipos de carne presentes en cada muestra. ....	48
Tabla 6: Comparativa coeficiente de correlación sobre novilla y cerdo a 580nm .....	65
Tabla 7: Comparativa correlación en carne de novilla a 1040nm .....	67
Tabla 8. Coeficientes de correlación máximos y sus respectivas longitudes de onda. Carne de Novilla .....	68
Tabla 9: Comparativa correlación en carne de cerdo a 762nm .....	70
Tabla 10 Coeficientes de correlación máximos y sus respectivas longitudes de onda. Carne de Novilla .....	71
Tabla 11: Comparativa conjunta PLS en carne de novilla y cerdo .....	81
Tabla 12: Comparativa novilla VIS-NIR promedio, MIP e INVMIP .....	86
Tabla 13: Comparativa novilla SWIR promedio, MIP e INVMIP .....	89
Tabla 14: Comparativa cerdo VISNIR promedio, MIP e INVMIP .....	92
Tabla 15: Comparativa cerdo SWIR promedio, MIP e INVMIP .....	94



# Parte I: Introducción



# Capítulo 1

## *Introducción y motivación del trabajo*

Este capítulo sirve como introducción de este trabajo y contextualiza la tecnología utilizada, la imagen hiperespectral, en el ámbito de la industria alimentaria y más concretamente en el análisis de productos cárnicos. Se comenzará estudiando los métodos que predominan actualmente en el análisis de la calidad y composición de los productos alimenticios para, posteriormente, explicar las ventajas y la reciente necesidad de implementar sistemas de imagen hiperespectral en este sector.

### **1.1 Análisis de control y calidad en la industria alimentaria**

En la industria alimentaria el análisis de la calidad de los productos es fundamental, debido a que la información que nos proporciona es útil tanto para evitar posibles problemas de salud provocados por el mal estado de los alimentos que ingerimos, como para conocer las propiedades que poseen dichos alimentos desde el punto de vista de la nutrición humana.

#### **Seguridad alimentaria**

Existe la posibilidad de que los consumidores puedan llegar a sufrir algún tipo de intoxicación debido a una contaminación de la materia durante el proceso de elaboración. Enfocándonos en los productos cárnicos, esta posibilidad es un riesgo debido a que un mismo producto proviene de una gran cantidad de materias primas diferentes y se comercializa sin los controles de calidad suficientemente exhaustivos para verificar y garantizar la calidad de todos los productos individualmente.

Por ello, sería necesario que la industria cárnica adoptase nuevas tecnologías que le permitan garantizar la calidad de la carne y dar información precisa de la composición en cada uno de los productos durante el proceso de elaboración.



Como consecuencia de la sociedad actual y el ritmo de vida que llevamos, la industria alimentaria estos últimos años ha centrado todos sus esfuerzos en conseguir mediante la ingeniería de la alimentación alargar la vida útil de los alimentos manteniendo la calidad a través de mejorar las técnicas de conservación. Es por eso por lo que los análisis de seguridad son tan importantes a lo largo de todas las fases del proceso de producción y deben garantizar que se mantiene una calidad alta de la carne y de sus características sensoriales.

El análisis de control de la calidad está enfocado principalmente en determinar la composición química de la carne, y a partir de ello poder identificar la presencia de aditivos u otros componentes. También se encarga de medir la evolución del estado de la carne cuando se somete a ciertos procesos naturales o provocados como por ejemplo observar el deterioro de las propiedades sensoriales con el transcurso del tiempo o sometiendo el producto a cambios de temperatura significativos.

Esos aditivos encontrados en el producto alimenticio pueden tener diferentes orígenes a lo largo de todo el proceso de elaboración: desde la materia prima a través de los pesticidas, por medio de sustancias liberadas por microorganismos localizados en la carne, cuando se añaden conservantes artificiales para alargar la vida útil del producto e incluso al final del proceso de producción al añadir químicos que mejoran ciertas propiedades sensoriales como la textura o el sabor. El poder identificar todas estas sustancias, y asegurar la presencia o ausencia de aditivos alimentarios, es algo muy valorado en el mercado, debido a que actualmente se está cuestionando el impacto que tiene en nuestra salud y las posibles enfermedades que puede provocar el uso de aditivos [1].

La normativa actual acerca de la seguridad alimentaria establece un método que únicamente obliga a realizar la trazabilidad de muestras por lotes, lo que no es un método de seguridad eficaz para evitar episodios como el de las “vacas locas” u otras crisis ocurridas en los últimos años en distintos sectores alimentarios.

Además, se ha descubierto que bacterias presentes en el ganado se están haciendo resistentes a tratamientos antibióticos debido a que en la propia alimentación de las reses se incluye antibiótico para favorecer el crecimiento y esto causa que las bacterias poli-resistentes hayan desarrollado un mecanismo de adaptación, lo que reduce la eficacia del medicamento. Esto supone un riesgo real para el consumidor que puede sufrir infecciones de carácter grave [2]. Poder identificar y detectar a tiempo antibiótico en la carne para prevenir su comercialización supone un nuevo reto para el control alimentario.

Por todo lo comentado anteriormente, la seguridad alimentaria ha adquirido tanta importancia que se han abierto líneas de investigación para mejorar y potenciar el desarrollo urgente de tecnologías más eficientes en este aspecto, capaces de garantizar la calidad de la carne en la misma línea de producción.

Actualmente la metodología de los análisis no se centra en el seguimiento de cada producto individualmente, sino como se ha mencionado previamente, de un lote de productos se realiza una única medida sobre un porcentaje pequeño de muestras y los resultados se extrapolan al conjunto





de productos que forman el lote. Este sistema no es el más fiable en cuanto a garantía de la calidad. Además, los análisis durante proceso de control se llevan a cabo en laboratorios externos y son procedimientos complejos, bastante tediosos y largos, muchas veces el tiempo necesario para obtener los resultados es mayor que el tiempo de vida del producto.

De lo anterior podemos detectar las siguientes necesidades en la seguridad de la industria alimentaria: la revisión de todos y cada uno de los productos elaborados y la importancia de hacerlo en tiempo de producción para que no lleguen al mercado productos sin el nivel de calidad requerido.

### **Información nutricional**

Todo lo referido a la seguridad alimentaria sumado al hecho de que recientemente, la población tiene una mayor conciencia y sentimiento de responsabilidad como consumidores, frente al efecto directo que tiene una buena alimentación sobre el estado de salud, hace que la trazabilidad actual y la información presente en las etiquetas sea escasa o no todo lo completa que podría ser.

Por trazabilidad se entiende todos los procedimientos que permiten conocer la procedencia de la materia prima o componentes, la ubicación a lo largo del proceso de elaboración y la trayectoria del producto o lote. Es el historial de los alimentos y una herramienta para la calidad y seguridad alimentaria. La información que se encuentra actualmente en las etiquetas no es individualizada del producto que se va a consumir, sino general de todo el lote del que proviene la carne. Las etiquetas nos dan la información básica acerca del alimento que compramos como la fecha de caducidad, el lote del que proviene, la cantidad de producto, los ingredientes, e información nutricional. Cuánto más completa y precisa sea esa información nutricional mejor podrán los consumidores realizar las elecciones en base a sus necesidades.

Por lo tanto, una información nutricional individual y detallada, que se ajuste de forma específica a cada producto, proporciona un gran avance en la industria alimentaria y será una gran ventaja a la hora de posicionarse en el mercado ya que son los consumidores los que demandan una mayor información de los productos en general, en este caso cárnicos, para estar más protegidos e informados tanto en el ámbito de la seguridad alimentaria como en el nutricional.

### **Resumen**

La industria alimentaria requiere dedicar recursos a la investigación para conseguir y mejorar los siguientes aspectos: detectar la presencia o ausencia de aditivos alimentarios en los productos elaborados; cuantificar la composición nutricional de cada uno de los productos pudiendo establecer niveles objetivos de los productos que se comercializan y garantizar la seguridad alimentaria evitando la salida al mercado de productos que no cumplan los requisitos mínimos, estén contaminados o contengan elementos no deseados.



## Técnicas convencionales para el análisis de control y calidad

En los apartados anteriores se ha reflejado la problemática general y las limitaciones de las técnicas utilizadas en los análisis de seguridad y calidad en la industria alimentaria. En este punto se incidirá con más detalle en la metodología de dichas técnicas, detectando las debilidades existentes y poniendo de manifiesto las posibles mejoras y avances que podrían llevarse a cabo a través de tecnologías como la espectroscopía de infrarrojo cercano, descrita en el siguiente apartado y en la que se basa este trabajo.

La calidad en la industria alimentaria se mide mediante un panel sensorial. Debido a que existe una gran cantidad de indicadores de calidad para los alimentos, es muy complicado cuantificarlos a través de parámetros físicos por el elevado número de variables por lo que este método no resulta ser muy objetivo. En lo que se refiere a propiedades cuantificables del producto, existen pocos sensores capaces de medir, y para los que existen, se necesita realizar ajustes constantes para realizar cada medida [3].

En cambio, la detección de contaminantes y aditivos artificiales que se encuentran en los alimentos sí puede ser cuantificada por procedimientos físico-químicos y, por ello, se pueden definir indicadores de seguridad y calidad relevantes. Es un proceso tedioso ya que el propio contenido de cada uno de los productos individualmente puede ser complejo, diferente y no uniforme. Las técnicas de control estándar más utilizadas en este ámbito son: el análisis infrarrojo y Raman [4][5][6], la fluorescencia [7], la espectroscopía UV-VIS [8], la espectroscopía atómica [9], la resonancia nuclear magnética [10], los tera-hercios [11], la espectroscopía dieléctrica [12], la cromatografía de líquidos y gases [13] y la espectrometría de masas [14].

Para todas estas técnicas, la metodología utilizada es la misma. Primero se obtiene una muestra, se transporta a un laboratorio externo y se analiza utilizando alguno de los procesos citados anteriormente. Es cierto que los resultados obtenidos son muy exactos, de gran precisión, pero son procedimientos muy complejos en los que las muestras analizadas requieren de una preparación previa y el tiempo para realizar el análisis es necesariamente elevado. Estas limitaciones imposibilitan aplicar estas técnicas en la línea de producción y poder analizar todas las muestras en tiempo real.

Actualmente los equipos situados en la línea de producción sólo están preparados para detectar contaminantes metálicos por medio de rayos X o para medir la grasa general en vez de realizar medidas más exactas y con mayor detalle para productos individuales. Esto se debe a que todas las técnicas de análisis requieren de una preparación previa y de un análisis complejo específico para cada tipo de producto, lo que implica estar haciendo ajustes continuamente para cada medida y eso no es viable en una línea de producción. Además, es difícil implantar un sistema que analice cada muestra en un tiempo reducido de análisis en la cadena de producción. El sistema estaría formado por los equipos convencionales utilizados en los laboratorios, para los que obtener resultados precisos requiere de mucho más tiempo.



## Espectroscopía infrarroja

Uno de los métodos más utilizados en los laboratorios para analizar la composición química de los alimentos es la espectroscopía infrarroja. Existen distintas técnicas que se basan en esta tecnología, como la Espectroscopía Infrarroja de Transformada de Fourier o la Espectroscopía Raman, muy utilizadas por obtener resultados altamente precisos. Son técnicas que requieren sistemas complejos, continua calibración y un proceso previo de preparación de la muestra, por lo que su uso se limita únicamente al laboratorio y no son técnicas implantables en la línea de producción.

### 1.2 Espectroscopía de infrarrojo cercano (NIRS) en la industria alimentaria

Como ya se ha comentado previamente, en la industria alimentaria existe la necesidad de analizar de forma precisa la calidad y la composición de la carne, sin afectar al producto y sin alterar el proceso de elaboración. Las técnicas que se emplean actualmente tienen grandes limitaciones. La espectroscopía de infrarrojo cercano, aunque también forma parte de la familia de técnicas mencionadas recientemente, posee la singularidad de poder realizar medidas corto periodo de tiempo y no necesita continuas calibraciones, ni de una preparación específica previa de las muestras.

Por lo tanto, las ventajas de esta tecnología son que se pueden realizar medidas casi instantáneas sin alterar la muestra porque es un método no destructivo, con sistemas de captura más sencillos, siendo más fácil de manejar e implantar en la cadena de producción. Esta tecnología en investigación muestra una gran potencialidad para caracterizar y cuantificar un gran número de parámetros y propiedades de las muestras [15][16][17]. En este trabajo se aplicará a productos cárnicos, más concretamente a las hamburguesas.

#### Fundamentos de la espectroscopía de infrarrojo cercano

La espectroscopía NIRS analiza el espectro en el infrarrojo cercano entre 700 y 2500nm, pudiendo extenderse también hasta el visible (400-700nm). Cuando se trabaja en el visible se habla de espectroscopía atómica y en el rango del espectro del infrarrojo cercano se trata de espectroscopía molecular.

La espectroscopía molecular consiste en analizar la absorción de los enlaces moleculares. Cuando se radia la molécula con la energía luminosa adecuada, ésta es absorbida por los enlaces moleculares provocando una serie de vibraciones moleculares fundamentales que se encuentran representadas en la Figura 1. Se producen sobretonos como consecuencia de existir múltiples niveles energéticos posibles para que se produzcan estas transferencias de energía, y ocurren a frecuencias dobles a

partir de la fundamental, ejemplo al doble de la fundamental encontraremos el primer sobretono,

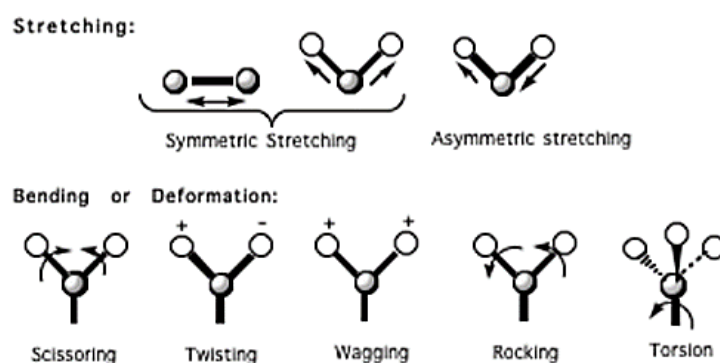


Figura 1. Vibraciones moleculares por radiación luminosa

al doble del primer sobretono estará el segundo sobretono y así sucesivamente. También podemos encontrar frecuencias de absorción características para los grupos funcionales, que son combinaciones de átomos típicos que dan lugar a moléculas mayores como O-H por ejemplo. Además, cuando se producen vibraciones fundamentales en distintas partes de la molécula se producen las combinaciones de bandas. Todas estas vibraciones son diferentes y por tanto específicas para cada molécula, lo que permite diferenciar entre compuestos y conocer con precisión la composición química de la muestra al analizar la información del espectro.

### Metodología de la espectroscopía de infrarrojo cercano

Las muestras se analizan a partir de dos fenómenos físicos, la transmisión difusa que necesita preparación de las muestras para que presenten un tamaño uniforme y con un grosor menos a 1 cm, y la reflexión difusa, que no necesita preparación y se puede utilizar para medir muestras de mayor grosor, con elevado esparcimiento o *scattering* de forma más rápida y sencilla. Los sistemas utilizados en este trabajo miden el segundo fenómeno.

Existen varios dispositivos para capturar los fenómenos anteriormente descritos a partir de los cuales se extrae la información, los más comunes son los monocromadores, los espectrómetros de array de diodos y los filtros interferenciales. Los filtros son los elementos más baratos por lo que se suelen implementar en el sistema final para abaratar su coste, una vez conocida la región del espectro de interés.

La limitación más característica de esta tecnología es la necesidad de crear un modelo de calibración sólido para cada una de las aplicaciones que se quieran desarrollar. Se debe entrenar el sistema analizando un gran número de muestras de todo tipo para poder realizar la calibración óptima específica y que factores como la humedad ambiente, el tamaño y la forma de la muestra, el polvo, etc. no alteren los resultados. Esto se debe a que lo que se obtiene es el espectro de cada muestra con información de múltiples propiedades. Cada muestra tiene una forma, humedad y composición diferente que da lugar a la obtención de un espectro diferente en cada medida. Por lo tanto, la

complejidad del proceso reside en conseguir diferenciar y extraer los parámetros relevantes de forma independiente de la muestra que se esté analizando.

Una vez diseñado un modelo de calibración sólido para la aplicación de interés, los parámetros cuantificables se pueden clasificar de la siguiente manera:

- **Composición química:** se puede conocer la composición química de la muestra, identificando de forma cualitativa las moléculas dependiendo de donde se sitúen los picos de absorción en el espectro, además de cuantificar la cantidad de esas moléculas en la muestra. Los parámetros más típicos en la carne son la proporción de grasa, proteína y humedad. Las medidas son muy susceptibles y se ven fácilmente afectados por factores que modifiquen el espectro. Los resultados se contrastan con medidas estandarizadas de referencia obtenidas de análisis químicos.
- **Características técnicas o tecnológicas:** características de la muestra como el pH, color, energía, firmeza, cambios después de procesos como congelación y cocinado, etc. No son muy precisos debido a la variabilidad de las muestras.
- **Atributos sensoriales:** estimaciones del sabor, textura, jugosidad, ternura. Como son criterios subjetivos, se intenta correlar la sensación de catadores con datos del espectro.

**Clasificación en categorías:** Se establecen criterios de clasificación a partir de muestras con características conocidas que se utilizan como entrenamiento, se necesita una gran cantidad para contemplar todas las posibles variaciones que puedan existir en las muestras no conocidas que se quieran clasificar. Posibles clasificaciones son, según la calidad, la cantidad de grasa, si la carne está cruda o congelada, si está adulterada, etc.

### 1.3 Objetivos de este trabajo

Como se ha mencionado en los puntos anteriores el objetivo general de este trabajo es desarrollar una aplicación enfocada al análisis de hamburguesas basada en la tecnología de espectroscopía de infrarrojo cercano de reflectancia o reflexión difusa (NIRS).

Se desea caracterizar y cuantificar la composición química de hamburguesas elaboradas con diferentes tipos de carne, centrándose fundamentalmente en establecer los porcentajes de grasa y proteína. Para ello se utilizarán dos sistemas de imagen hiperespectral (HSI, *Hyperspectral Imaging*) de espejo rotativo, que trabajan en diferentes rangos del espectro, en el visible y parte del infrarrojo cercano (VIS-NIR) y el otro sistema en el infrarrojo cercano de onda corta (SWIR, *Short Wave Infrared*). Ambos sistemas han sido desarrollados por el Grupo de Ingeniería Fotónica del Departamento TEISA de la Universidad de Cantabria.

Se realizarán comparaciones entre los resultados obtenidos con ambos sistemas, y los diferentes tipos de carne, de novilla y de cerdo, para estudiar en qué casos la aplicación es más precisa de forma cualitativa al clasificar los porcentajes de grasa o de carne de las muestras analizadas, contrastando los datos obtenidos con medidas de referencia.



Se tomarán medidas también de mezclas ya conocidas, para entrenar el sistema con el objetivo de conseguir cuantitativamente una mejor calibración en la obtención de los porcentajes de cada parámetro.

Para obtener y visualizar los resultados deseados hay que procesar la imagen hiperespectral captada por los sistemas HSI a través de diferentes técnicas de procesado. Otro de los objetivos es establecer la relación entre los resultados obtenidos, el tipo de carne analizada y las técnicas de procesado que se hayan aplicado. Se estudiarán y compararán las diferentes técnicas de procesado para establecer cuáles son más eficientes en cada caso.

Simplificando, los objetivos específicos que se desea alcanzar son los siguientes:

1. Clasificar cualitativa y cuantitativamente los porcentajes de grasa y proteína de las diferentes hamburguesas.
2. Clasificar hamburguesas en función de sus propiedades comparando los distintos tipos de carne.
3. Establecer cuáles son las técnicas de procesado y los rangos espectrales más adecuados para caracterizar cada componente.
4. Establecer qué sistema es más eficiente para medir los distintos tipos de carne y obtener una mejor información de los parámetros estudiados.

## 1.4 Estructura del documento

Es documento se estructura en 5 partes, divididas a su vez en capítulos con varios apartados cada uno:

- **Parte I: Introducción.** Contiene un único capítulo llamado “Introducción y motivación del trabajo”, es la primera parte del documento y sirve para contextualizar la temática, fijar los objetivos y definir la estructura del mismo.
- **Parte II: Teoría y estado del arte.** Contiene dos capítulos que proporcionan las bases teóricas necesarias para entender el funcionamiento de los sistemas aplicadas al análisis de hamburguesas.
  - o **Capítulo 2: Propiedades espectrales de la carne.** Descripción de los componentes principales de la carne y los porcentajes de dichos elementos presentes en las hamburguesas. Explicación de las propiedades ópticas de los tejidos y estudio del comportamiento espectral de los cromóforos que componen las hamburguesas.
  - o **Capítulo 3: Sistemas espectroscópicos.** Base teórica de imagen espectral y diferencia con imagen convencional RGB. Estudio del estado del arte de las diferentes técnicas de adquisición de imagen hiperespectral que utilizan los sistemas espectroscópicos de infrarrojo cercano más comunes.



- **Parte III: Materiales y métodos.** Consta también de dos capítulos centrados en la descripción de los equipos, de los tipos de carne utilizados y de las técnicas de procesado implementadas.
  - o **Capítulo 4: Materiales.** Descripción de las características de los sistemas utilizados en este trabajo (VIS-NIR y NIR)
  - o **Capítulo 5: Métodos. Técnicas de procesado espectral.** Descripción de las técnicas de análisis espectral empleadas en el desarrollo y ejecución del proyecto.
- **Parte IV: Resultados.** En dos capítulos se discutirán los resultados de clasificación cuantitativa y cualitativa del análisis de hamburguesas como ejemplo de producto cárnico.
  - o **Capítulo 6. Clasificación cuantitativa.** Resultados de la implementación de técnicas de cuantificación de los componentes constitutivos de las hamburguesas.
  - o **Capítulo 7. Clasificación cualitativa.** Resultados de la implementación de técnicas de discriminación cualitativa de las diferentes hamburguesas.
- **Parte V: Conclusiones finales y líneas futuras.** Capítulo final que resume las conclusiones del trabajo y las posibles líneas de acción en un futuro próximo.



## Parte II: Teoría y estado del arte



## Capítulo 2

### *Propiedades espectrales de la carne*

Se podría definir la carne como la parte comestible que proviene de los animales post mortem. Hay una gran variedad de especies de las que se obtienen productos cárnicos que se utilizan para el consumo humano. En este trabajo se estudian dos de los tipos de carne más comunes y que más se consumen como son la carne de vaca y de cerdo. En el ganado vacuno y porcino el porcentaje de músculo aprovechable para la producción de carne está en torno al 35%. Los componentes principales son grasa, proteína y agua que conforman entre el 97 y el 99% de la carne. El porcentaje restante, del 1 al 3% lo forman otros componentes como carbohidratos, vitaminas, hormonas, etc. [18].

- **Proteína.** La fracción proteica abarca el 20% del músculo aproximadamente y está compuesta por varios tipos como la proteína miofibrilar y proteína del tejido conectivo.
- **Grasa.** Dependiendo de cada animal el contenido de grasa puede variar y suponer entre el 1 y el 80% del peso del músculo. Los principales componentes son lípidos, ácidos grasos, fosfolípidos, acilglicéridos, etc.
- **Humedad.** Referida al contenido de agua en los músculos, las moléculas de agua en la carne pueden encontrarse en diferentes estados, dependiendo de las uniones de los enlaces de hidrógeno, causado por ejemplo por factores como la temperatura.

### 2.1 Fabricación de hamburguesas

La cantidad presente de cada uno de los tres componentes mayoritarios (agua, grasa, proteína) en la carne no es independiente, existe una alta relación con las proporciones existentes de los otros dos componentes [19]. En cambio, si se mezclan con otros ingredientes o sustancias adicionales como especias, almidón o agua para obtener otros productos cárnicos como por ejemplo salchichas o hamburguesas el coeficiente de inter-correlación entre los porcentajes de grasa, humedad y proteína es menor.

Este trabajo se centrará en el estudio de hamburguesas. Para ello, se fabricarán hamburguesas con distinta composición química, analizando dos tipos de carne diferentes y variando el porcentaje de



grasa en cada muestra, con el fin de entrenar el sistema y poder caracterizar y estimar el porcentaje de grasa de todas las mezclas de las hamburguesas fabricadas que se encuentren a la venta.

En el mercado existe una gran variedad de hamburguesas con distinta calidad y valor nutricional, esto se debe a que son un producto cárnico que se puede elaborar con distintas proporciones de sus componentes.

Según la Organización de Consumidores y Usuarios (OCU), en una hamburguesa de carne vacuno el porcentaje de grasa debería estar en torno al 6% para tener un valor nutricional óptimo, pero la media ronda el 15% de grasa, existiendo algunos casos donde llega al 25% lo que conlleva un menor contenido en proteínas. El porcentaje medio de contenido en proteína, cuando la grasa no alcanza valores muy altos, se sitúa alrededor del 15%. La hamburguesa, como la carne en general, no contiene casi hidratos de carbono, y los que aparecen provienen del almidón usado al realizar la mezcla o de los azúcares añadidos [20].

Teniendo en cuenta esos datos se puede fabricar una hamburguesa variando la cantidad de componentes en función de la proporción que se quiera conseguir. Un ejemplo de fabricación de hamburguesa de vacuno es el siguiente, del total del producto 70 grs son de carne picada de vacuno, 10 grs de grasa, 1.5 grs de sal, 0.5 grs de pimienta negra y 10 grs de agua [21].

## 2.2 Propiedades ópticas de los tejidos

Determinar la cantidad presente en la carne de los tres componentes principales que la forman, es decir conocer el porcentaje de grasa, humedad y proteína se puede considerar uno de los puntos del análisis del proceso de producción más básicos e importantes, debido a que conocer la composición química de la materia empleada en la fabricación de un producto puede ayudar a optimizar la composición del producto final.

Para poder caracterizar hamburguesas es necesario conocer las propiedades ópticas de los tejidos en general, es decir la forma en que la luz interactúa con la materia. Esto nos permitirá entender los principios en los que se fundamentan los sistemas HSI utilizados en este trabajo para realizar las medidas. Cada tejido interactúa de forma diferente con luz en función de su composición química, según la distribución y la cantidad de cromóforos existentes. Esto se debe a la capacidad de la luz de penetrar en los tejidos, reaccionar con la materia y devolver información de esa interacción que tras su detección nos permite saber que componentes ha interrogado.

Conocer las propiedades ópticas de la muestra que queremos analizar, en este caso carne, más concretamente hamburguesas de novilla y de cerdo nos facilita saber que sistemas y que técnicas de procesamiento de imágenes nos interesa utilizar para obtener la información observando el rango del espectro de mayor interés. Se miden a través de varios parámetros, los más significativos son el coeficiente de absorción  $\mu_a$  ( $\text{cm}^{-1}$ ), el coeficiente y la función de esparcimiento scattering  $\mu_s$  ( $\text{cm}^{-1}$ ) y  $p(\theta, \varphi)$  ( $\text{sr}^{-1}$ ) respectivamente, y el índice de refracción del tejido  $n'$ .

## Coeficiente de Scattering

El esparcimiento óptico o scattering es el esparcimiento que sufre la luz al chocar con partículas cuyo índice de refracción es diferente al del medio dónde se encuentra, o cuando la luz atraviesa un medio en el cual el índice de refracción sufre cambios continuos.

El primer caso se modela con la teoría de scattering de Mie, la cual aproxima las partículas del medio que provocan el esparcimiento como esferas perfectas. Tratando todas las partículas que contribuyen al esparcimiento como esferas, se puede estimar el comportamiento del esparcimiento de la luz en el conjunto de tejido analizado. Cuando las partículas tienen un tamaño comparable (similar o mayor), a la longitud de onda de la luz se denomina scattering Mie. Por otro lado, el scattering Rayleigh es un caso concreto del anterior aplicado a partículas mucho más pequeñas que la longitud de onda.

El coeficiente de scattering reducido  $\mu'_s$ , en función de la longitud de onda  $\lambda$ , es un parámetro que sirve para caracterizar el comportamiento de tejidos cuando incide luz sobre ellos y se puede modelar mediante la siguiente ecuación [22]:

$$\mu'_s(\lambda) = a' \left( f_{Ray} \left( \frac{\lambda}{500(nm)} \right)^{-4} + (1 - f_{Ray}) \left( \frac{\lambda}{500(nm)} \right)^{-b_{Mie}} \right),$$

donde  $a'$  es un factor que escala el término dependiente a la longitud de onda,  $f_{Ray}$  es la fracción del Rayleigh scattering y  $b_{Mie}$  es la potencia de scattering. En la ecuación la longitud de onda a la que se quiere conocer el coeficiente reducido de scattering está normalizada en base a una longitud de referencia de 500 nm y está descrita en función de la contribución del Rayleigh scattering por un lado y el scattering Mie por otro.

La medida del coeficiente de scattering es difícil de realizar teniendo en cuenta que se realiza midiendo la transmitancia de la luz a través de un tejido ( $T_c$ ), de grosor  $L$ . Si el grosor de la muestra es demasiado grande, no podrá llevarse a cabo la medida.

$$\mu_s = -\frac{\ln(T_c)}{L}$$

Por ello, los tejidos analizados requieren una preparación previa de la muestra y en ocasiones conseguir el grosor tan pequeño necesario para obtener resultados válidos es una tarea complicada (del orden de 100  $\mu m$  normalmente). Además, si la mezcla tiene una heterogeneidad alta, también dificulta la medida del coeficiente de scattering.

## Coefficiente de Absorción

Los tejidos en mayor o menor medida absorben parte de la radiación que incide sobre ellos a través de las vibraciones moleculares de sus componentes. Se definen como cromóforos a los componentes de la muestra que absorben luz en una región espectral de interés. Cada cromóforo tiene su propio espectro de absorción para cada longitud de onda. Por tanto, el coeficiente de absorción depende de la composición química del tejido y del grosor o recorrido de la luz a través del material. El coeficiente de absorción se puede definir con la siguiente expresión:

$$\mu_a = -\frac{1}{T} \frac{\delta T}{\delta L}$$

Donde  $T$  es la fracción de variación de luz restante, la que no absorbe al recorrer un medio durante un variación o incremento de la trayectoria de la luz a través de dicho medio  $L$  (cm). Esta expresión de la variación de la luz transmitida permite observar que, a medida que la trayectoria recorrida es mayor, la intensidad de la luz decrece de manera exponencial debido a la absorción producida.

$$T = e^{-\mu_a L} = 10^{-\epsilon C L} = e^{-4\pi n'' L / \lambda}$$

En la expresión anterior se muestra la absorción en función de varios factores. Se puede relacionar el coeficiente de absorción con la concentración de cromóforos, cuando se expresa la intensidad de la luz transmitida a través de un medio idealmente sin esparcimiento donde se haya un cromóforo cuya concentración es  $C$  (mol  $L^{-1}$ ) y  $\epsilon$  es su coeficiente de extinción ( $cm^{-1} M^{-1}$ ). La intensidad de transmisión de la luz, y por tanto el coeficiente de absorción del cromóforo, también puede expresarse en función de  $n''$  que es el índice imaginario de refracción del medio.

$$T = e^{-\mu_a L} = e^{-4\pi n'' L / \lambda} \quad \rightarrow \quad \mu_a = 4\pi n'' / \lambda$$

El coeficiente de absorción de un tejido se conforma por medio de la suma de las  $i$  diferentes contribuciones de absorción de cada uno de los cromóforos presentes en la muestra de tejido [22].

$$\mu_a = \ln(10) \sum_i C_i \epsilon_i$$

El estudio realizado en este trabajo se centrará en el análisis del coeficiente de absorción de los diferentes cromóforos más característicos presentes en la carne, y se realizará en las longitudes del espectro donde la absorción del cromóforo sea más significativa pudiendo así detectar y cuantificar los componentes químicos que conforman la muestra.

Los cromóforos dominantes en el cálculo del coeficiente de absorción de un tejido generalmente son el agua y la sangre (hemoglobina). En la sangre podemos analizar la contribución de la hemoglobina oxigenada ( $HbO_2$ ) y la hemoglobina desoxigenada ( $Hb$ ), en el caso de la carne es mayor la contribución de la hemoglobina oxigenada en la absorción ya que no se encuentra sangre circulando al ser un producto fabricado post mortem. La absorción producida por el agua varía dependiendo del estado en el que se encuentre en el tejido, y esto depende de los enlaces moleculares.

En la Figura 2. se observa el coeficiente de absorción de los diferentes componentes de un tejido genérico en el rango del espectro de 300 a 1200 nm, donde se puede apreciar el comportamiento de algunos de los cromóforos también presentes en la carne.

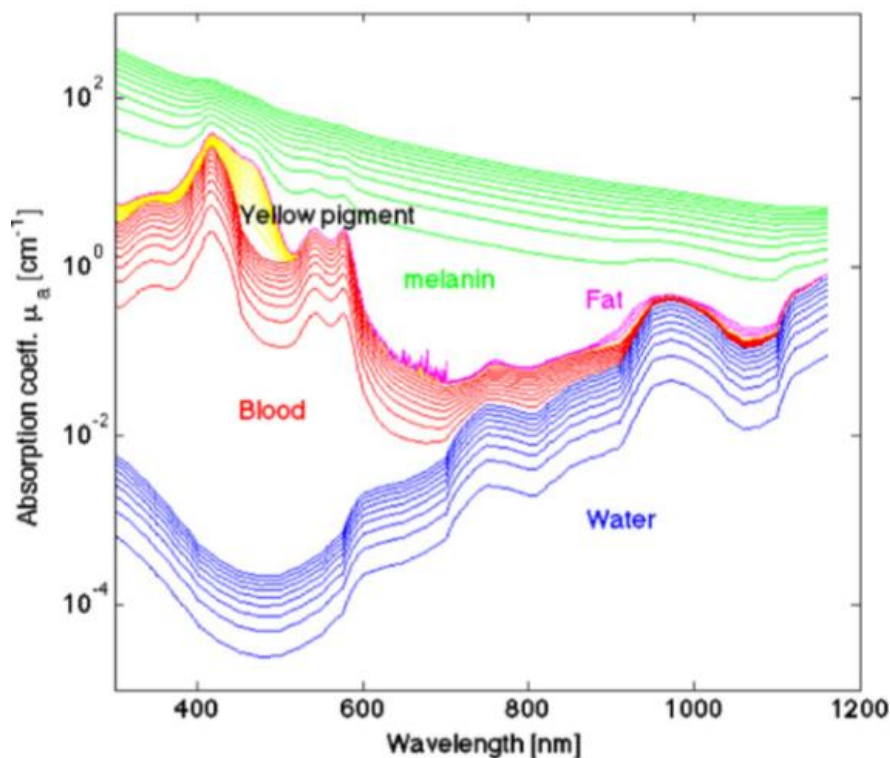


Figura 2: Coeficientes de absorción de los componentes de un tejido genérico en el rango del espectro de 300 a 1200 nm [22]

Para la caracterización de la carne, interesa cuantificar también otros cromóforos como la grasa cuyo pico dominante de absorción se encuentra a 930nm [22], proteína, lípidos, colágeno, etc. Pudiendo conocer así la composición química, y por tanto, las propiedades nutricionales de cada muestra individualmente.

### 2.3 Picos de absorción de los principales cromóforos de la carne

Es necesario conocer cuáles son los picos de absorción más significativos de los cromóforos presentes en la carne y en qué longitudes de onda se manifiestan para poder correlar esa información con el espectro obtenido tras realizar las medidas de las hamburguesas pudiendo así contrastar resultados.

En este trabajo se estudiará el comportamiento espectral de los principales cromóforos de la carne en dos regiones del espectro diferentes, en VIS-NIR, que abarca el visible y parte del infrarrojo cercano (400-1000 nm aproximadamente) y la región SWIR, zona de onda corta del infrarrojo (comprendida entre 1000 y 2000 nm aproximadamente).

Es interesante estudiar el comportamiento de cada uno de los componentes que forman la carne y cómo contribuyen al coeficiente de absorción en las respectivas longitudes de onda en ambas bandas espectrales, debido a que hay cromóforos que presentan resultados más destacables en términos de absorción en una única una región, como por ejemplo sucede con el agua, los lípidos y el colágeno, cuyas ventanas de absorción de mayor relevancia se encuentran en la región SWIR. Conocer el coeficiente de absorción en las longitudes de onda significativas, es el primer paso a la hora de cuantificar la cantidad del correspondiente cromóforo presente en la muestra.

## La hemoglobina

Es el componente de la sangre cuya absorción es más notable. Esto se debe a que presenta una concentración de 150g por litro de sangre lo que le permite transportar más oxígeno que ningún otro componente presente en la sangre como por ejemplo el plasma. Conforman alrededor del 95% de la masa de las células rojas de la sangre.

Dependiendo de los enlaces moleculares se distingue entre hemoglobina oxigenada ( $\text{HbO}_2$ ), cuando se enlaza con el oxígeno y hemoglobina desoxigenada cuando se manifiesta sin enlaces moleculares, Hb. En el estudio de productos cárnicos los picos de absorción dominantes se deben a la hemoglobina oxigenada debido a que no hay sangre circulando.

Los picos más notables se sitúan en la región VIS-NIR como se puede observar en la Figura 3. En el caso de la hemoglobina oxigenada los picos de absorción se encuentran en torno a 344, 414, 542 y 576 nm. y la absorción de la hemoglobina desoxigenada se manifiesta en las longitudes de onda de 431 y 556 nm.

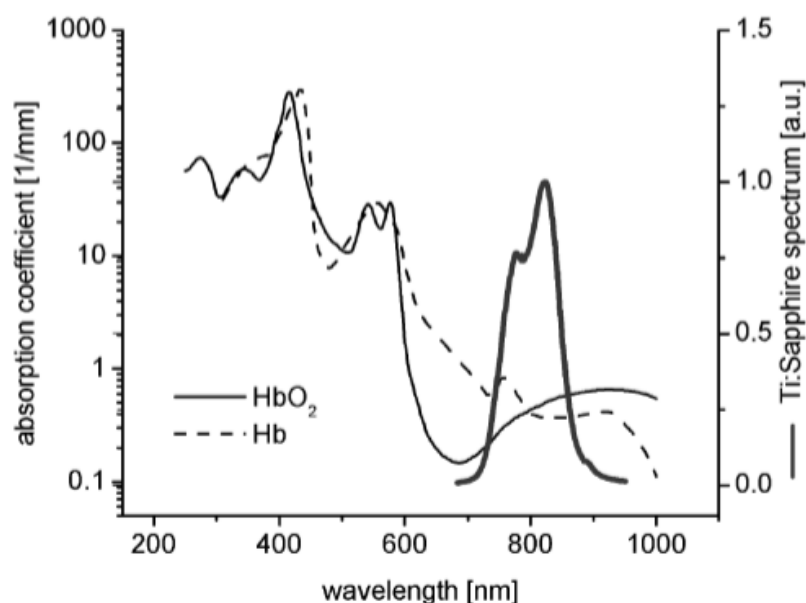


Figura 3: Picos de absorción de la hemoglobina [23]



## Humedad ( $H_2O$ )

En la primera parte del capítulo 2 se menciona la humedad como uno de los tres componentes principales de la carne. Ahora el objetivo es determinar en qué zonas del espectro es mayor la ventana de absorción para este cromóforo.

Los picos de absorción del agua en las longitudes de onda de 970 y 1180 nm se deben a las vibraciones producidas por un sobretono del enlace molecular O-H, el primer sobretono del movimiento molecular de estiramiento o *stretching* del enlace O-H produce un pico de absorción alrededor de los 1430 nm; en 1930 nm se encuentra otro pico debido a una combinación de estiramiento o *stretching* y deformación y por último, a 1975 nm se produce la absorción debido al segundo sobretono de una deformación por curvamiento o *bend* del enlace O-H [24].

Como se ha comentado anteriormente, los coeficientes de absorción más altos se encuentran en la región SWIR y, por lo tanto, los que más peso tienen a la hora de determinar el coeficiente de absorción total debido al agua. El pico de 1440nm tiene un coeficiente de absorción 60 veces mayor que el pico en 970 nm y en 1940nm es en torno a 260 veces mayor en comparación con el pico de 970 nm. Esto sirve para entender la magnitud de la diferencia existente entre los valores de absorción en la región VIS-NIR frente a la región SWIR.

## Lípidos

Los lípidos forman parte del tejido adiposo de los animales por ello al inicio del capítulo 2 se engloban dentro de la grasa. Los picos de absorción de los lípidos ocurren debido a las diferentes vibraciones moleculares del enlace C-H en las longitudes de onda: 920, 1040, 1210, 1730 y 1760 nm. En torno a 1430 nm el pico de absorción se debe al primer sobretono del movimiento molecular de *stretching* del enlace O-H.

Distintos tipos de lípidos como colesterol, triglicéridos y fosfolípidos tienen espectros muy parecidos entre ellos y como ocurre también en el caso del agua, los picos más distintivos de cada uno de ellos, con mayor coeficiente de absorción se producen en la región SWIR. Puede incluso llegar a ser 80 veces mayor comparando la absorción a 1760 nm frente a la que se produce a 920nm. Esto significa que la sensibilidad a la hora de determinar la cantidad de lípidos y agua presentes en la muestra sería mucho mayor si se mide en la región SWIR que en la VIS-NIR [24].

## Colágeno

El colágeno es una proteína estructural que se agrupa en forma de fibras. Está presente en el tejido conjuntivo, óseo y cartilaginoso. Es un componente clave, abundante tanto en los tejidos blandos como duros. La absorción propia de este cromóforo se produce en las siguientes longitudes de onda: 930, 1000, 1200, 1500, y en torno a 1700 nm. Igual que sucede en los dos casos anteriores, y como se puede observar en la Figura 4. los principales picos de absorción se sitúan en la región SWIR.



Si los picos que presentan un coeficiente de absorción mayor se comparan con los de menor, resulta que la absorción es 5 veces mayor a 1200 nm que a 1000nm e incluso puede llegar a ser 38 veces mayor en 1500 nm respecto a la absorción en 930 nm [25]

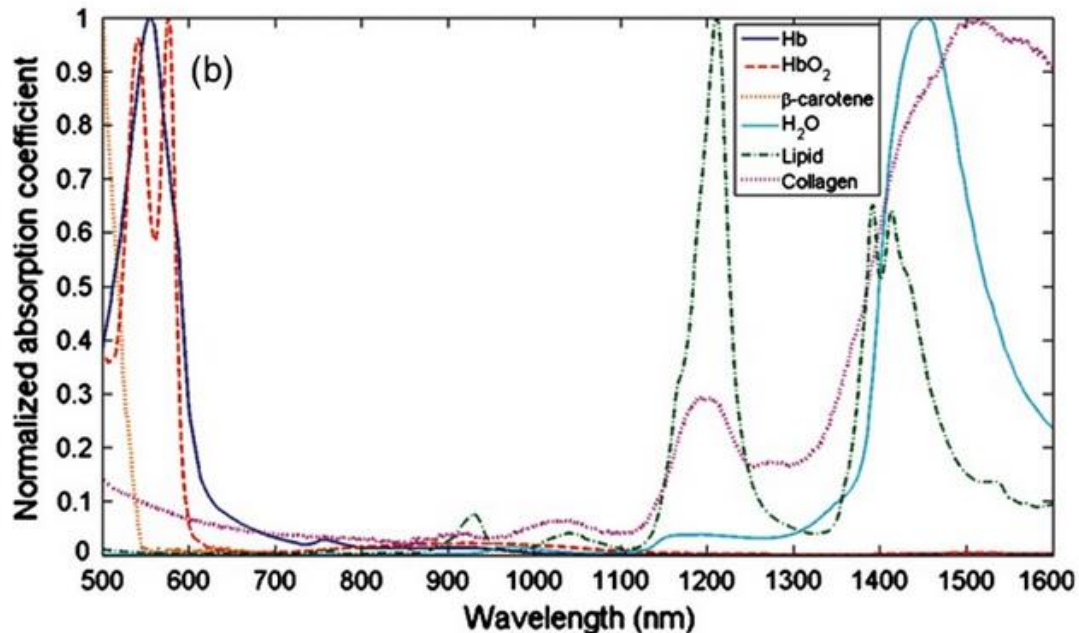


Figura 4: Coeficiente de absorción normalizado del colágeno, lípidos, hemoglobina y agua [24]

## Proteína

Es uno de los componentes fundamentales de la carne y del tejido animal en general. Es la base de estructuras del cuerpo como la piel y los músculos, está compuesta por aminoácidos. También tiene una función importante en la generación de enzimas como anticuerpos. Por tanto, se considera un elemento fundamental para que el cuerpo y tejido vivo animal funcionen correctamente.

Los picos de absorción característicos de la proteína se producen por los sobretonos de la vibración de moléculas N-H y se localizan alrededor de las longitudes de onda: 1187, 1510, 1690 y 2265 nm. Todos ellos en la región SWIR, situándose algunos en el límite superior de esta, en torno a los 2000 nm [26].

En la mayoría de las muestras es difícil diferenciar la absorción provocada solamente por proteína debido a su cercanía a otros picos de absorción de la grasa o del agua, cuyas ventanas de absorción son mayores al igual que los coeficientes de absorción como se puede observar en la Figura 5, por lo que la proteína resulta indistinguible, ya que su aportación en esas longitudes de onda pasa desapercibida. Es por ello por lo que para la caracterización de la proteína se obtienen mejores resultados cuando se analizan muestras deshidratadas o con bajo contenido en agua [27].



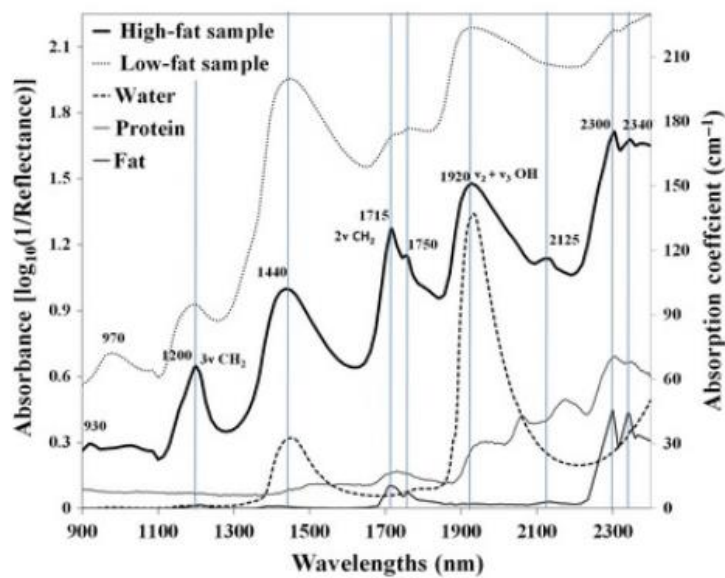


Figura 5: Coeficientes de absorción de la proteína [27]

## Grasa de cerdo

El último componente por estudiar es la grasa de cerdo, aunque ya se ha analizado la absorción en los lípidos, conocer la absorción de esta grasa en concreto puede facilitar la comprensión de los resultados obtenidos en el trabajo, debido a que la variación en el porcentaje de grasa de las muestras se consigue añadiendo grasa de cerdo.

Los picos de absorción de la grasa de cerdo en la región VIS-NIR, como se puede observar en la Figura 6, se sitúan en torno a las longitudes de onda: 762, 830, 930 y 1040 nm [28]. Si se compara con la absorción de los lípidos se observa que tienen un comportamiento similar sobre todo en 930 y 1040 nm.

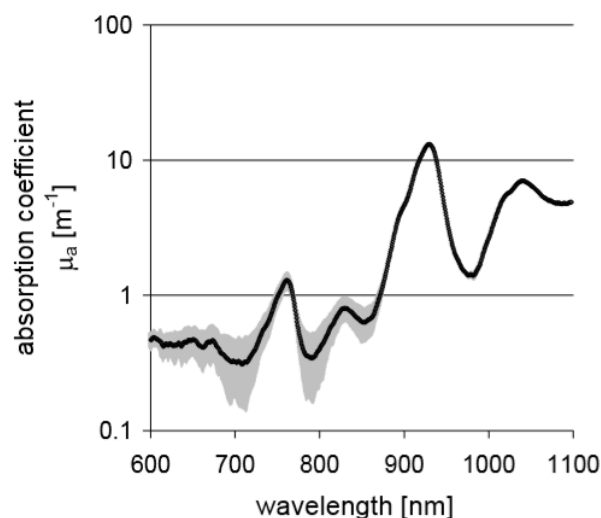


Figura 6: Coeficientes de absorción de la grasa de cerdo [28]

# Capítulo 3

## *Sistemas espectroscópicos*

### 3.1 Sistemas HSI

Los sistemas utilizados en este trabajo son sistemas espectroscópicos HSI, los cuales se basan en detectar la reflectancia o la transmitancia difusas proveniente de la muestra correctamente iluminada por una fuente de luz blanca. La información capturada se almacena en imágenes hiperespectrales, que tras aplicar diferentes técnicas de procesado nos permiten conocer la composición química de la materia examinada.

#### **Imagen espectral**

Una imagen espectral es una imagen que proporciona una gran resolución espectral además de la propia resolución espacial. Cuántas más longitudes de onda contenga, mayor será la información del espectro que se pueda analizar y, por tanto, mayor será el número de parámetros de la muestra que se podrán caracterizar.

Las imágenes espectrales, generan una gran cantidad de datos, por cada punto del espacio se almacena la información de la reflectancia que contiene cada una de las longitudes de onda analizadas, por lo que el procesado de la matriz tridimensional resultante es más complejo que en otras técnicas de imagen.

La imagen hiperespectral es la técnica de imagen espectral que más información es capaz de recoger. Al igual que la imagen multiespectral, puede proporcionar resolución espectral también fuera de la región visible, en cualquier rango del espectro (300-2600 nm), como en el infrarrojo y dentro de la región ultravioleta [29]. En definitiva, ambas son técnicas de imagen espectral ya que pueden obtener información de múltiples longitudes de onda y en el caso de la imagen hiperespectral es capaz de almacenar información de una mayor parte del espectro.

La espectroscopía convencional, como se aprecia en la Figura 7.a. no puede proporcionar información espacial precisa. Permite conocer el comportamiento de la muestra a determinada longitud de onda, pero no relacionarlo con la región de la muestra dónde sucede, perdiendo con ello la referencia espacial. Sin embargo, con las imágenes hiperespectrales esa información no se pierde ya que la referencia espacial también está almacenada en la matriz tridimensional, aunque al representar el espectro no se puedan relacionar ambos parámetros simultáneamente.

## Imagen RGB y diferencias con imagen espectral

Por otro lado, una imagen monocromática solo contiene información espacial para una determinada longitud de onda. Una imagen RGB, en cambio, contiene dentro de la región visible, información espacial solo en 3 longitudes de onda concretas correspondientes a los 3 colores primarios, azul (435 nm), verde (545 nm) y rojo (630 nm). Es por ello por lo que las imágenes RGB permiten obtener información espectral en tres puntos, al tiempo que toda la información espacial, ya que fuera de esas longitudes de onda no detecta ningún componente.

La principal diferencia entre las imágenes espectrales y RGB se puede apreciar mejor en la Figura 7.b. donde se toma un pixel y se compara la curva de reflectancia en esa posición de la imagen hiperspectral y la intensidad de la imagen RGB en ese mismo punto. La gráfica de la intensidad proporciona información en tres longitudes de onda concretas, mientras que la curva de reflectancia contiene información de la región del espectro observada de forma continua.

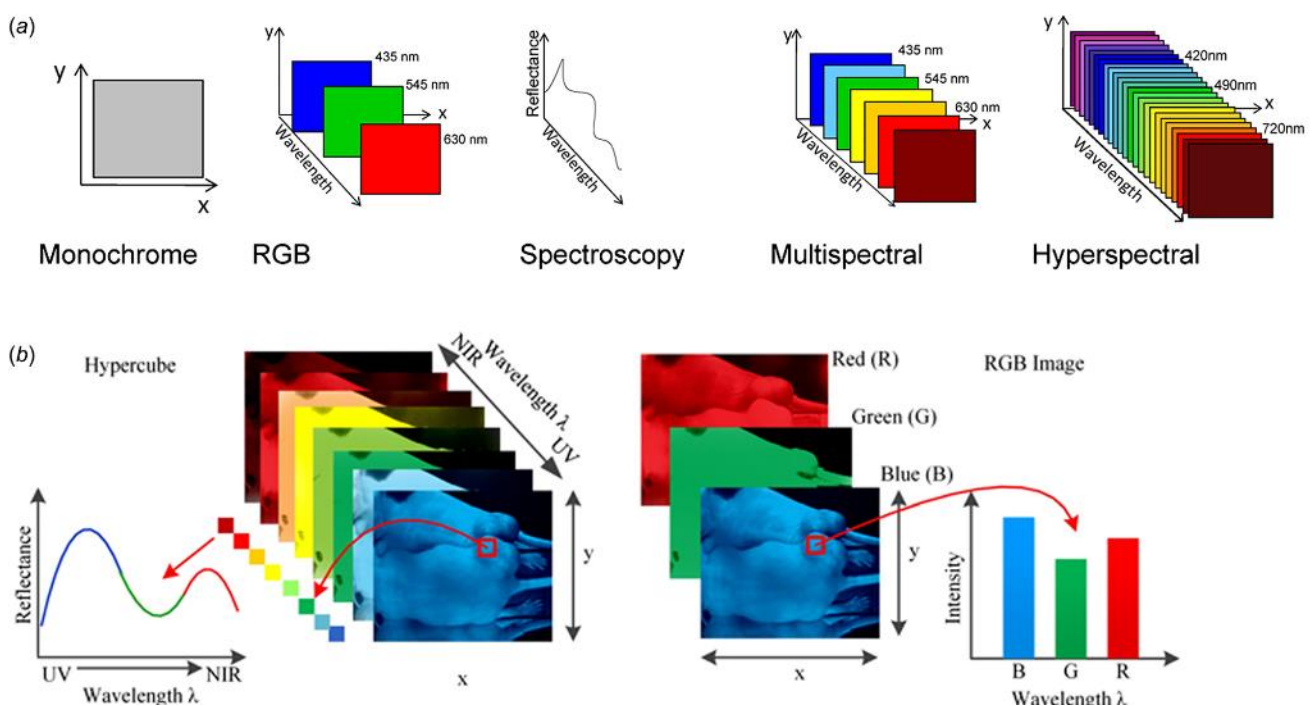


Figura 7: (a) Diferencias entre imagen monocromática, RGB, multispectral e hiperspectral. Representación gráfica de espectroscopía. (b) Diferencias entre la representación espectral de la imagen hiperspectral y la intensidad de una imagen RGB. [29]

## Métodos de obtención de imagen hiperspectral

Existe un gran porcentaje de los sistemas utilizados a la hora de obtener imágenes hiperspectrales. Todos ellos no pueden medir la dimensión espacial de la muestra al completo y obtener la información espectral en una única captura. Necesitan mecanismos o dispositivos que permitan realizar un barrido e ir adquiriendo la información por partes.

Hay varios métodos para llevar a cabo este escaneo de la muestra, adquirir los datos y conformar la imagen hiperespectral.

**Non-scanning.** Técnica que no realiza ningún escaneo, también denominada *snapshot*. Se obtiene la imagen hiperespectral de toda la muestra en una sola captura. Sistemas muy costosos o con poca resolución espacial.

**Spectral Scanning.** Se realiza un escaneo espectral, es decir se captura una imagen bidimensional por cada longitud de onda que se vaya a medir, por tanto, la imagen hiperespectral se conformará almacenando cada una de las imágenes monocromáticas escaneadas. Este barrido se realiza mediante filtros ajustables, dependiendo de la longitud de onda de interés se utilizarán filtros paso bajo o paso banda.

**Spatial Scanning.** Se realiza un escaneo espacial para cada posición de la muestra que el sistema quiere analizar. Se obtiene una imagen bidimensional con toda la información espectral de un vector  $x$  que recorre la muestra línea por línea  $(x, \lambda)$ . Este barrido se puede implementar de diferente forma [30]:

- *Pushbroom*: se realiza el escaneo de la muestra, línea por línea, por medio de algún dispositivo mecánico que mueva la muestra o el sistema de detección.
- *Whiskbroom*: se escanea cada línea de la muestra a través de espejos que se mueven de forma mecánica de manera que para posición del espejo se obtiene la información espectral de un único pixel de la línea que se va interrogando. Se realiza un escaneo línea por línea y dentro de cada línea, a su vez, un barrido punto por punto. Es un método de escaneo más lento, pero se consigue una alta resolución espacial y espectral.

**Spatiospectral scanning.** Se realiza el escaneo combinando los dos métodos anteriores, se realiza el barrido para una posición concreta  $x$  y cierta longitud de onda.

**Rotary mirror scanning.** Sistema de escaneado que utiliza un espejo rotativo y un subsistema óptico con el objetivo de leer el espectro de una línea de la muestra en cada posición del espejo. Es una combinación de las dos técnicas de escaneado espacial, ya que se produce un barrido línea por línea de la muestra al igual que usando *pushbroom*, gracias a un espejo rotativo (como se utiliza en la técnica *whiskbroom*) pero usando un subsistema óptico que permite barrer varios píxeles a la vez.

Los elementos dispersivos como se muestra en la Figura 8. para el escaneo espacial tanto para *whiskbroom*, *pushbroom* y para el caso del no escaneo o *snapshot*, son prismas, rejillas y combinación de prismas y rejillas, respectivamente. En cambio, para el escaneo espectral se utilizan filtros ajustables como se ha comentado anteriormente.

El rango del espectro donde son aplicables los distintos métodos es amplio menos el caso del escaneado espectral, siendo medio debido a que se ve limitado por el uso de los filtros. La resolución espectral es alta para las técnicas de escaneado espacial, véase la Tabla 1, mientras que para el

escaneado espectral y *snapshot* puede llegar a ser baja, para mejorar esta propiedad se necesitaría usar equipos más costosos.

Modalidad	Resolución Espacial	Resolución Espectral
Imagen Instantánea (Snapshot Imaging)	Dado por la resolución del sensor, $\sim 100 \times 100$ px/ resolución espacial depende de la óptica	Dado por la óptica, 50-100ch./ $\sim 10$ nm
Sistema de Filtros (Staring Systems)	Dado por la resolución del sensor, $\sim$ Mpx	6-24ch./ $\sim 100$ nm (rueda de filros) variable/ $\sim 10$ -40 nm (LCTF)
Sistema de Barrido Puntuales (Whiskbroom)	Dado por el dispositivo mecánico (ambos x e y)	Dado por el detector y espectrógrafo, 100-1000 ch./ $\sim 2$ -10 nm
Sistemas de Barrido de línea (Pushbroom)	Dado por la cámara (x) y el motor (y) de resolución lineal/ $\sim$ Mpx	Dado por el detector y espectrógrafo, 100-1000 ch./ $\sim 2$ -10 nm
Escáner de Espejo rotativo (Rotating Mirror Scanner)	Dado por la cámara (x) y el motor (y) de resolución angular/ $\sim$ Mpx	Dado por el detector y espectrógrafo, 100-1000 ch./ $\sim 2$ -10 nm

**Tabla 1: Diferencias entre las técnicas de adquisición de imagen hiperespectral [31]**

También se ha de tener en cuenta que los sistemas escaneado espacial y de escaneado espectral son de imagen hiperespectral mientras que en el caso de *snapshot* son sistemas de imagen multiespectral. Por ello al usar sistemas *whiskbroom* y *pushbroom*, o una combinación de ellos como en el caso del escáner de espejo rotativo, el número de datos que se puede generar y recoger en la matriz tridimensional es mayor que utilizando el resto de los métodos y su procesado será más complejo y lento. Al contrario que sucede con los sistemas *snapshot*, cuyo procesado es más rápido y menos complejo al recoger menos cantidad de datos aunque a costa de perder resolución espectral [29].

Con relación al coste de implementar un sistema u otro se ha de tener en cuenta que los sistemas basados en técnicas de escaneo espacial son los menos costosos, debido a que los filtros del escaneado espectral y el sistema de *snapshot* son equipos más complejos de implementar y por tanto económicamente más costosos. Dependiendo del objetivo que queramos conseguir y la resolución de los parámetros que se quieran estudiar, teniendo en cuenta las propiedades de los diferentes sistemas previamente mencionadas, se valorará la utilización de un sistema u otro.

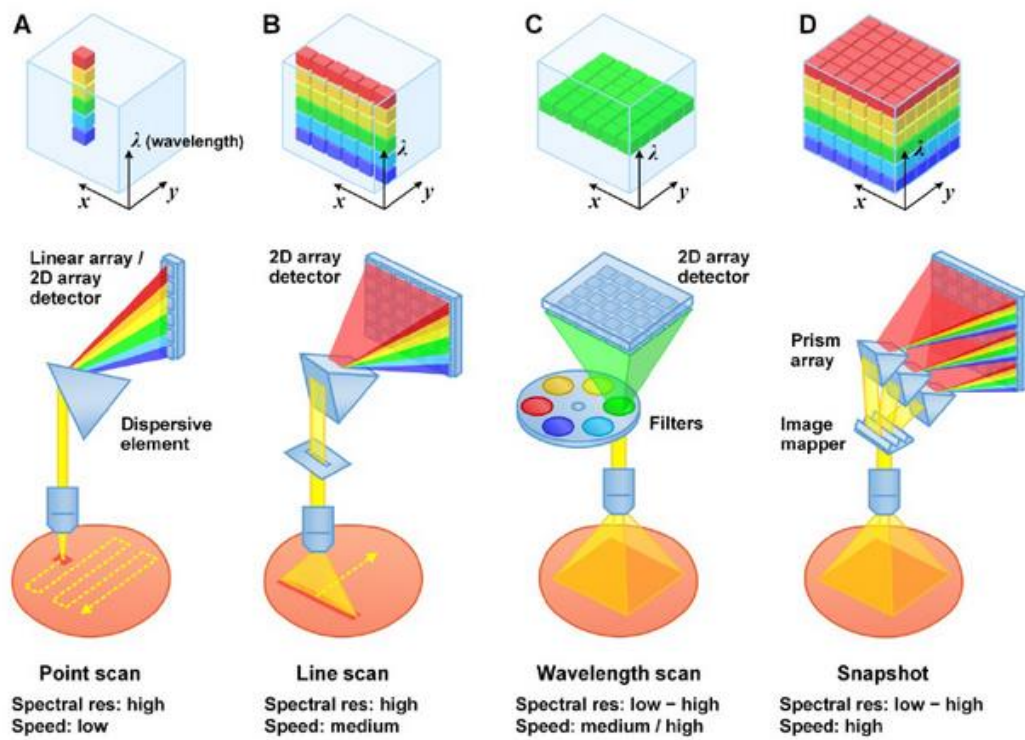


Figura 8: Funcionamiento de las técnicas de adquisición de imagen hiperespectral [32]



## Parte III: Materiales y métodos

# Capítulo 4

## *Materiales*

Este capítulo consta de dos partes, la primera se centra en describir más detalladamente los sistemas HSI que se utilizan en este trabajo. Se trata de conocer las características y propiedades ópticas de los componentes que forman el sistema a través de su funcionamiento y sus especificaciones técnicas. La segunda parte de este capítulo explica el proceso inicial de medidas al que se sometió a las muestras analizadas.

### 4.1 Sistema HSI de espejo rotativo

Para la realización de medidas llevadas a cabo en este trabajo se utilizan dos sistemas HSI que emplean la misma técnica de adquisición de la imagen hiperespectral, un escáner mediante un espejo rotativo (Figura 9). Ambos sistemas se diferencian en el rango espectral de trabajo, es decir, en el rango del espectro que pueden capturar. Un sistema trabaja en la región VIS-NIR (longitudes de onda entre 400 y 900 nm) y otro equipo es capaz de capturar en la región SWIR, zona de onda corta del infrarrojo cercano (longitudes de onda entre 1000 y 1700 nm).

El prototipo en el que se basa el funcionamiento de ambos sistemas de espejo rotativo, y el cual se utiliza para realizar las medidas en la región VIS-NIR del espectro en este estudio, se desarrolló en el Trabajo de Máster de José Alberto Gutiérrez-Gutiérrez [33] y ha sido publicado en el artículo [31].



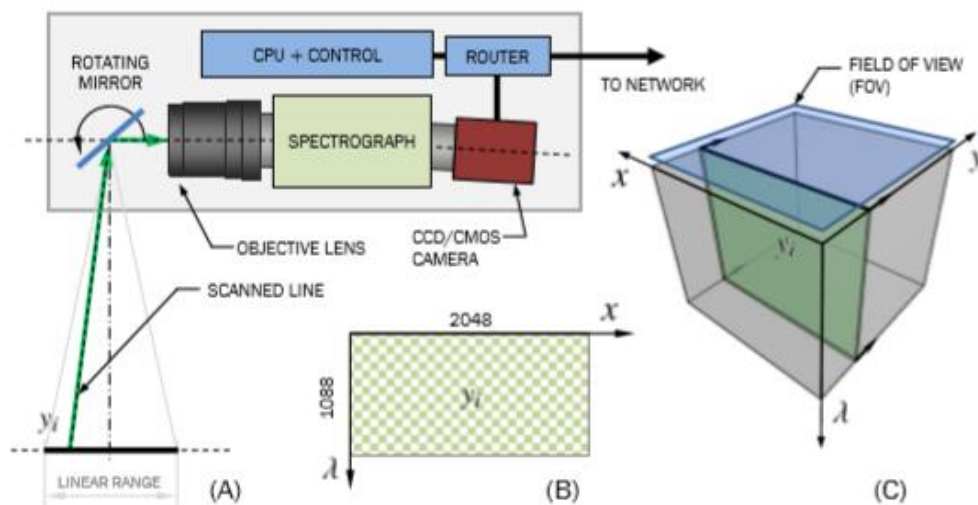


Figura 9: Esquema del funcionamiento de los sistemas HSI de espejo rotativo [31]

La ventaja del sistema de espejo rotativo es que permite realizar un barrido y obtener información de la muestra en su totalidad sin necesidad de mover el sistema de detección que realiza la medida o la muestra. Esto lo convierte en un equipo compacto y portable, que proporciona mayores facilidades a la hora de tomar medidas respecto a los sistemas dependientes de dispositivos mecánicos externos para realizar el barrido de la muestra y, por tanto, presenta una mayor adaptabilidad a diferentes entornos de medida.

Los diferentes componentes que forman este sistema de imagen hiperespectral con espejo rotativo pueden agruparse dentro de los siguientes grandes bloques: subsistema óptico; subsistema actuador, encargado de rotar el espejo para cambiar la posición de la línea de la muestra que se está capturando; subsistema de control, interfaz que nos permite manejar el sistema y subsistema de comunicaciones que permite conectar el sistema a una red de comunicaciones. [31]

### Principales componentes ópticos del sistema HSI

Los componentes ópticos fundamentales del sistema HSI son la cámara y subsistema óptico formado por el espectrógrafo y la óptica de enfoque que permiten capturar la información espectral de una línea de píxeles de la muestra. Ambos sistemas (VIS-NIR y SWIR) siguen el mismo esquema y método de funcionamiento, su diseño es análogo sólo se diferencian en las especificaciones técnicas tanto de la óptica, de la cámara y del espectrógrafo.

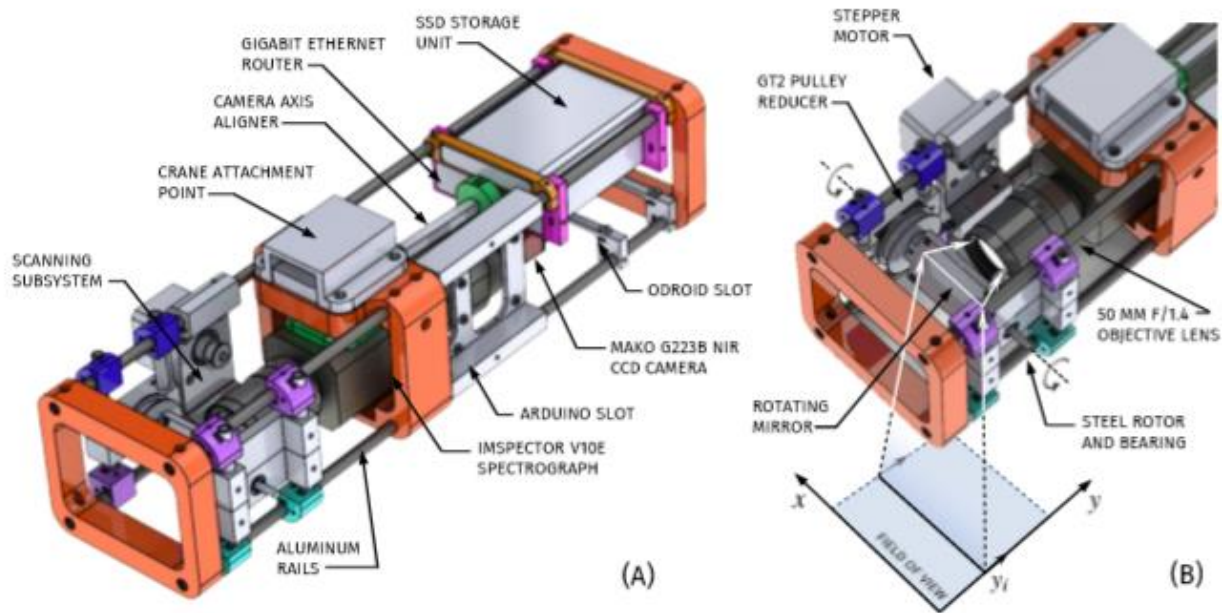


Figura 10: Esquema general del sistema. (A) Componentes principales. (B) Sistema de escáner del espejo rotativo. [31]

### Óptica dispersiva y de enfoque

En las siguientes tablas se especifican las propiedades técnicas de la óptica de enfoque para cada uno de los sistemas.

- Sistema HSI rango VIS-NIR:

Óptica de enfoque	Manual con zoom
Distancia focal	5-50 mm
Apertura	Hasta f/1.4
Acoplamiento	Montura C del ImSpector

Tabla 2: Especificaciones técnicas óptica de enfoque sistema HSI rango VISNIR [33]

## Cámara

- Sistema HSI rango VIS-NIR

El sistema que captura en el rango VIS-NIR del espectro utiliza una cámara de tecnología CCD. El modelo usado es AVT Mako G 223B NIR, funciona con la interfaz GigE Vision cuyas especificaciones se encuentran en la Tabla. El acceso para el control de la cámara se tiene que llevar a cabo a través de Ethernet. [33]

Interfaz	GigE
Mega Pixels	2.0
Resolución	2048 x 1088
FPS	50
Tipo de sensor	CMOS Global Shutter
Sensor	CMV2000-NIR
Pixel Pitch ( $\mu\text{m}$ )	5,5 x 5,5
$L \times W \times H$ (mm)	60,5 x 29 x 29
Rango de temperaturas	5 a 45 °C
Auto Iris	None
Montura	C

Tabla 3: Especificaciones técnicas cámara AVT Mako G 223B NIR [33]

- Sistema HSI rango SWIR:

Sistema HSI rango SWIR pendiente de publicación, el modelo de la cámara utilizada es Xenics XEVA 883 (InGaAs) XC130. Longitud de onda entre 900nm y 1700nm.

## Espectrógrafo

Es el dispositivo óptico capaz de capturar una línea de píxeles sobre la superficie de la muestra y, para cada punto sobre la línea del material, obtiene su espectro combinando la información espacial y espectral en una imagen que se proyecta sobre sensor CMOS, al descomponer las longitudes de onda con un sistema óptico de lentes, prismas y rejilla (*grating*) holográfica.

En ambos sistemas el espectrógrafo utilizado emplea un elemento PGP (*Prism-Grating-Prism*). En la Figura 11 se observa el esquema que sigue este tipo de espectrógrafos que utiliza componentes ópticos como prismas, combinados con rejillas, para corregir la aberración óptica que introduce el elemento PGP.

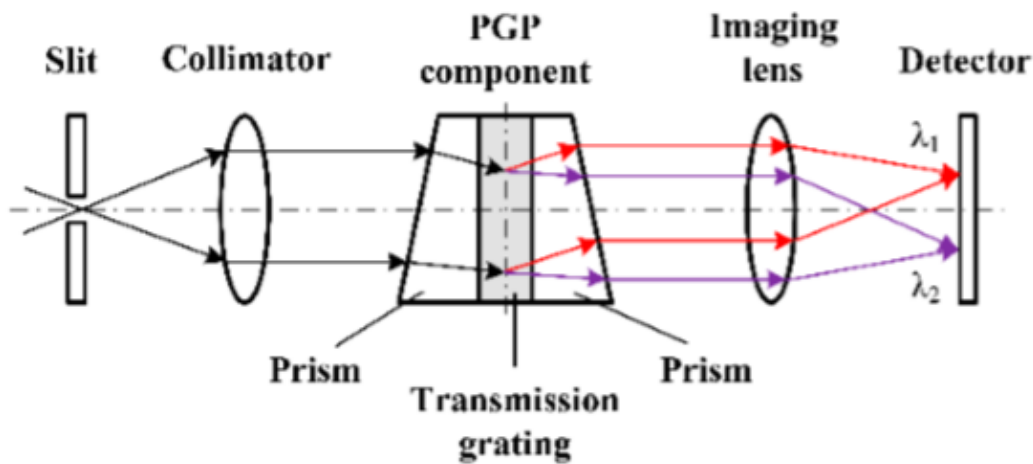


Figura 11: Esquema de un espectrógrafo con un componente PGP [35]

- Sistema HSI rango VIS-NIR:

En la Tabla 6. se pueden observar las principales características del espectrógrafo Speciem ImSpector V10E utilizado en el sistema que trabaja en la región VIS-NIR del espectro.

Rango espectral	400-1000 nm
Resolución espectral	2.8 nm (with 30µm slit)
Resolución espacial	rms spot radius 9 µm
Dispersión	97.5 nm/mm
Smile	1.5 µm
Keystone	1 µm
Tamaño de imagen (espectral x espacial)	Max 6.15 x 14.2 mm
F / #	F/2.4
Anchura de ranura	30 µm (13, 18, 50, 80 and 150 µm)
Tamaño efectivo de ranura	14.2mm
Eficiencia total (Típica Stray Light)	50% Independientemente de la polarización
Stray Light	0.5% (halogen lamp, 633nm notch filter)
OEM Size	(W) 60 x (H) x 60 x 175mm
Peso	1100g
Montura de las lentes	Standard C-mount adapter
Montura de la cámara	Standard C-mount adapter

Tabla 4: Especificaciones técnicas técnicas espectrógrafo sistema HSI rango VISNIR [33]

- Sistema HSI rango SWIR:

Sistema HSI rango SWIR pendiente de publicación, el espectógrafo utiliza una lente Navitar.

## 4.2 Preparación de las muestras

En este trabajo las medidas realizadas se hacen sobre mezclas de hamburguesas de dos tipos de carne diferentes, de novilla y de cerdo. En las mezclas se varía el porcentaje de grasa añadiendo tocino a la carne, para tener cuantificado el porcentaje de grasa que debería detectar el sistema HSI. Cada una de las mezclas se dispone en una de las cavidades de la cubeta. Hay que tener en cuenta que la capacidad máxima de cada cavidad no permite más de 2 gr de mezcla, en el mejor de los casos, ya que los dos tipos de carne y la grasa tienen diferente densidad. Al ser cantidades tan pequeñas las que entran en cada cubeta y tener que estar presente porcentajes de carne y porcentajes de grasa tan específicos, resulta inviable conseguir mezclas con los porcentajes deseados.



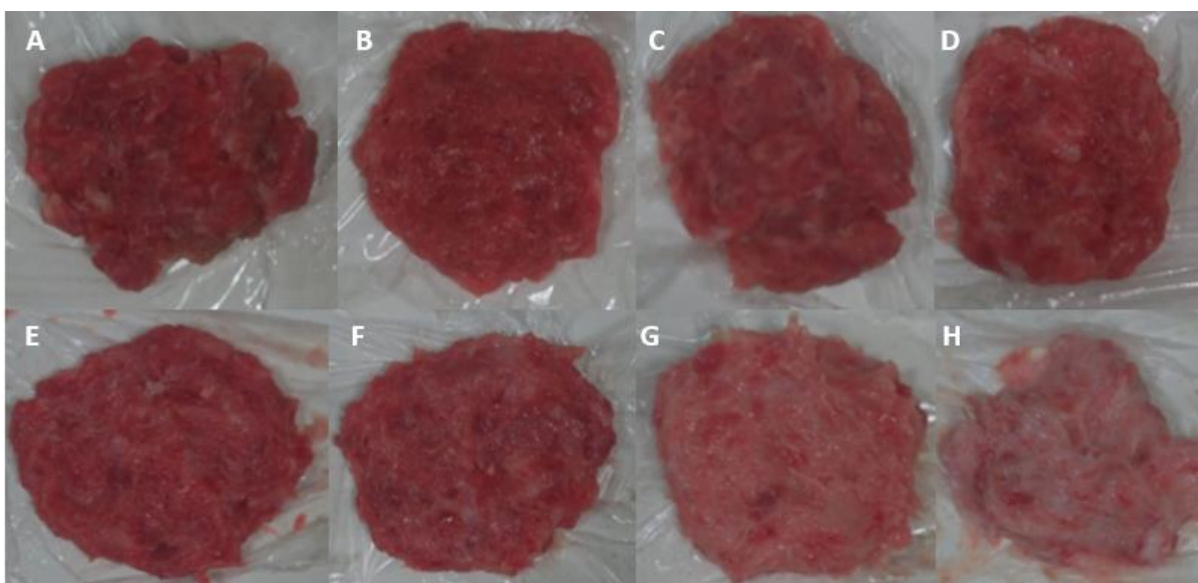
Figura 12: Cantidad de carne de novilla (izquierda) y tocino (derecha) utilizados en la elaboración de la muestra de novilla con un 75% de carne

Por tanto, el método a seguir será crear mezclas de 10gr donde resulte más fácil conseguir las proporciones necesarias de grasa y de carne. La dificultad reside en que la mezcla debe ser lo más homogénea, para que los 2gr que se depositaran en la cubeta contengan la misma proporción que la mezcla de 10gr; y compacta para evitar que la superficie sea demasiado granulada y minimizar así posibles efectos de esparcimiento/*scattering* que podrían afectar o distorsionar los resultados de las medidas. Esa compactación se hace manualmente, amansando la mezcla lo máximo posible.



**Figura 13: Proceso de colocación de las muestras con distintos porcentajes de grasa en la cubeta**

En la Figura 14 se aprecia visualmente la variación de las proporciones de grasa y carne en las hamburguesas realizadas con carne de cerdo. A mayor porcentaje de grasa más claro es el tono de la mezcla resultante.



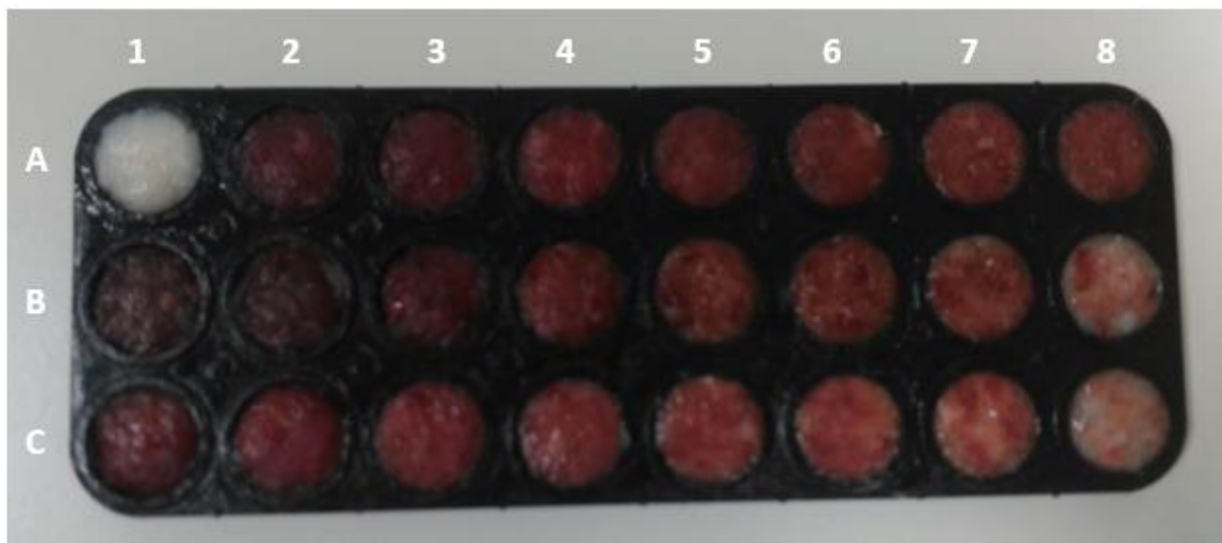
**Figura 14: Variación del porcentaje de grasa en las muestras de carne de cerdo. A 100% de carne, B 95% de carne, C 90% de carne, D 85% de carne, E 80% de carne, F 75% de carne, G 50% de carne y H 25% de carne**



La carne de novilla contiene de por sí menos grasa que la carne de cerdo, es por ello que, aun añadiendo la misma proporción de grasa a la carne inicial, la carne de novilla presenta un color más oscuro. Esta diferencia se puede apreciar en las Figuras 13 y 15.

### Composición mezclas de hamburguesa.

En este apartado se especifica la composición de todas las mezclas que se van a medir y su localización dentro de la cubeta. La cubeta, como se aprecia en la siguiente figura, consta de 24 cavidades repartidas en 3 filas de 8 columnas.



**Figura 15:** Distribución de las distintas mezclas en la cubeta tras efectuar las medidas. Comparar con la Figura 13 para observar la diferencia en el estado de la carne antes y después de la realización de las medidas.

La primera columna contiene la grasa de referencia, es decir, tocino al 100% y la carne de novilla y de cerdo también al 100%, ya que no contienen grasas añadidas aparte de las que contiene de por sí la carne. La segunda fila contiene mezclas de carne de novilla combinadas con grasa de tocino en una proporción ascendente de grasa partiendo del 100% de novilla y 0% de grasa, al 25% de novilla y 75% de grasa. La tercera fila sigue el mismo procedimiento que la segunda, pero en este caso con la carne de cerdo.

La variación de las proporciones de las mezclas queda reflejada en la siguiente tabla, según la distribución que siguen las cubetas, mencionada anteriormente.



<b>A1</b> Tocino (T) 100%	<b>A2</b> N 75% C 25%	<b>A3</b> N 50% C 50%	<b>A4</b> N 25% C 75%	<b>A5</b> A3 95% T 5%	<b>A6</b> A3 90% T 10%	<b>A7</b> A3 85% T 15%	<b>A8</b> A3 80% T 20%
<b>B1</b> Novilla (N) 100%	<b>B2</b> N 95% T 5%	<b>B3</b> N 90% T 10%	<b>B4</b> N 85% T 15%	<b>B5</b> N 80% T 20%	<b>B6</b> N 75% T 25%	<b>B7</b> N 50% T 50%	<b>B8</b> N 25% T 75%
<b>C1</b> Cerdo (C) 100%	<b>C2</b> C 95% T 5%	<b>C3</b> C 90% T 10%	<b>C4</b> C 85% T 15%	<b>C5</b> C 80% T 20%	<b>C6</b> C 75% T 25%	<b>C7</b> C 50% T 50%	<b>C8</b> C 25% T 75%

Tabla 5: Porcentajes de grasa y de los diferentes tipos de carne presentes en cada muestra.

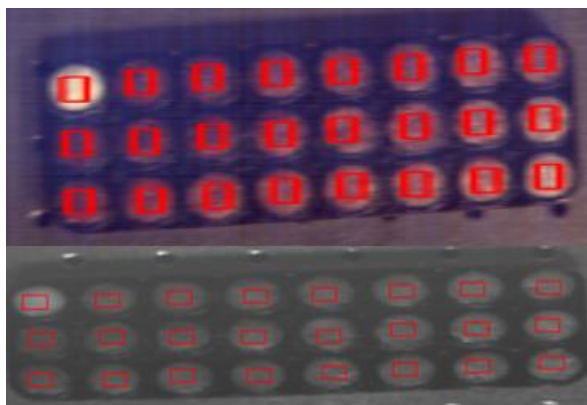
De cara a las técnicas de procesado para cuantificar el porcentaje de grasa de cualquier hamburguesa que se mida, se entrena al sistema con una variación de la grasa más pequeña dentro del rango de variación de la cantidad de grasa que presentan las hamburguesas en el mercado, que suele estar entre un 5% y un 25% aproximadamente, siendo el valor normal en torno al 13%, como se comenta en el capítulo 2. Cinco de las ocho mezclas preparadas tienen valores de grasa dentro de este rango, con el objetivo de realizar una clasificación del porcentaje de grasa de la forma más precisa posible.



## Capítulo 5

### *Métodos*

Antes de explicar las técnicas de procesamiento espectral, es necesario tener en cuenta que el software de los sistemas, tras realizar la medida, es capaz de seleccionar las regiones de interés (ROIs, *Regions of Interest*), es decir, zonas recortadas de la muestra, donde posteriormente se aplicarán las técnicas de procesamiento. En la Figura 16, se muestran las regiones seleccionadas por cada sistema. La parte superior se corresponde al sistema VIS-NIR y la inferior al sistema SWIR. Para facilitar el tratamiento de los datos espectrales se aplican ROIs con las mismas dimensiones en cada sistema.



**Figura 16: Distribución ROIs.(Superior) Sistema VIS-NIR. (Inferior) Sistema SWIR**

El software de los sistemas también realiza la calibración espectral [31]. La calibración espectral es necesaria en cualquier sistema de imagen hiperespectral debido a que la medida de reflectancia captada de una muestra contiene una combinación de espectros del objeto a medir, de los elementos ópticos del sistema y de la fuente de luz. Por eso, para poder distinguir la aportación del espectro causada únicamente por la muestra hay que calibrar el sistema espectralmente con un material con reflectancia difusa estandarizada (blanco)[34]. La ecuación para calibrar espectralmente la medida realizada para cada longitud de onda es la siguiente.

$$R(x, y, \lambda) = \frac{S_{mp}(x, y, \lambda) - B(x, y, \lambda)}{W(x, y, \lambda) - B(x, y, \lambda)}$$

donde  $B(x,y,\lambda)$  es la respuesta del sistema al negro o condiciones de oscuridad;  $S_{mp}(x,y,\lambda)$  es la medida de la muestra por el sistema y  $W(x,y,\lambda)$  es la medida del material con reflectancia estandarizada. De este modo, la reflectancia después de la calibración toma valores entre 0 y 1 ya que la medida de la muestra como máximo tendrá el mismo nivel de reflectancia que la referencia del material estandarizado.

## 5.1 Técnicas de procesado

El siguiente paso consiste en el procesado de los espectros obtenidos mediante diferentes técnicas y modelos matemáticos. Este paso es clave para la correcta determinación de la concentración de presencia de compuestos que constituye el núcleo principal de interés en este trabajo. Los métodos que se van a utilizar posibilitan la extracción de las características significativas de cada componente, grasa y carne. Para una correcta extracción de la información de interés, partiendo de los espectros de las hamburguesas y sus componentes, se necesita investigar qué métodos se van a desarrollar.

La cuantificación de presencia de parámetros sigue generalmente dos líneas: cualitativa y cuantitativa. Las técnicas de procesado implementadas en este trabajo se dividen según el mismo criterio, las técnicas aplicadas en la clasificación cuantitativa o las técnicas aplicadas a la clasificación cualitativa.

### 5.1.1 Técnicas de procesado para clasificación cuantitativa

En este trabajo se aplican dos métodos diferentes [36], se implementa un modelo de clasificación clásico de mínimos cuadrados (CLS, *Classical Least Squares*) que realiza un ajuste lineal de los valores medidos en cada longitud de onda. El segundo modelo se basa en procesos estadísticos de vectores propios. A partir de las medidas de las muestras se aplica el método de regresión PLS (*Partial Least Squares*) para obtener el modelo que describa la composición de las hamburguesas como carne o grasa.

#### *Classical Least Squares*

Este método basa el cálculo en una relación entre la respuesta espectral  $A_\lambda$ , el coeficiente de absorción  $\epsilon_\lambda$ , la trayectoria recorrida por la luz  $b$  y la concentración del componente de la muestra  $C$ . Si se considera el recorrido de la luz constante para todas las muestras, esta relación se puede expresar como [36]:

$$A_\lambda = K_\lambda C$$

donde el coeficiente de absorción y la trayectoria de la luz se combinan mediante la constante  $K_\lambda$ . Esta ecuación se resuelve midiendo la absorción de una muestra de concentración conocida, pudiendo obtener así el valor de  $K_\lambda$  a esa determinada longitud de onda. Conocido ese  $K_\lambda$ , se puede estimar la concentración de cualquier muestra en ese punto del espectro.

Para que la aproximación sea más exacta y no se vea esa estimación afectada por el ruido, errores de instrumentación o de fabricación de las muestras, es mejor aplicarla sobre varias medidas con concentraciones conocidas, intentando así ajustar linealmente la mejor solución para todas. El proceso descrito aproxima la concentración de un componente a una longitud de onda. Por tanto, para cada longitud de onda hay que estimar la concentración de los diferentes componentes a través de la respuesta espectral y de la absorción que generan. El cálculo queda reflejado con la siguiente ecuación:

$$\begin{bmatrix} A_{\lambda 1} \\ A_{\lambda 2} \end{bmatrix} = \begin{bmatrix} K_{a,\lambda 1} & K_{b,\lambda 1} \\ K_{a,\lambda 2} & K_{b,\lambda 2} \end{bmatrix} \begin{bmatrix} C_a \\ C_b \end{bmatrix} + \begin{bmatrix} E_{\lambda 1} \\ E_{\lambda 2} \end{bmatrix}$$

Expresada matricialmente obtendríamos la siguiente expresión donde la  $E_{\lambda i}$  representa el error del ajuste lineal que se comete al estimar las concentraciones.

$$A = KC + E$$

Una vez que la anterior ecuación se resuelve y la matriz  $K$  es conocida se pueden aproximar las concentraciones de cualquier muestra. Para poder resolver la ecuación se debe aplicar el procedimiento descrito en la ecuación (12), proceso necesario para que el cálculo de la inversa de la matriz siempre tenga solución.

$$K = AC'(CC')^{-1} \quad (12)$$

La mayor limitación de este método es que las aproximaciones lineales deben tener en cuenta y estimar el valor para todos los componentes de la mezcla, evitando así que algún componente no considerado interfiera en el análisis.

### ***PLS (Partial Least Squares)***

PLS es una técnica de descomposición espectral [36]. Durante el proceso de descomposición, PLS utiliza la información de las concentraciones de cada espectro, en lugar de descomponer primero la matriz espectral en un conjunto de vectores propios y puntuaciones, y retroceder en contra de las concentraciones como un paso separado. El método PLS analiza la matriz espectral para determinar cuáles son los vectores propios en función de las concentraciones de cada espectro. Los vectores propios son los espectros de las observaciones con características significativas a partir de los cuales se puede estimar las concentraciones del resto. La idea principal del PLS es conseguir la mayor cantidad de información de la concentración en los primeros vectores de carga. Simplemente tiene en cuenta la relación de correlación que existe entre los datos espectrales y las concentraciones que lo constituyen. Los datos espectrales y los datos de concentración pueden ser descompuestos en sus variaciones más comunes.

PLS es un proceso de un único paso, realiza la descomposición de los datos espectrales y de concentración simultáneamente. Como cada nuevo factor es calculado para el modelo, los valores se intercambian antes de que se elimine la contribución del factor de los datos sin procesar. Al igual que las matrices de datos reducidas, se usan ahora para calcular el siguiente factor, y el proceso se repite hasta que se calcule el número de factores deseados. La ventaja de PLS es que los vectores espectrales resultantes están directamente relacionados con los componentes de interés.

Se pueden implementar dos algoritmos diferentes, PLS-1 y PLS-2 [36]. Las diferencias entre ambos métodos son sutiles, sin embargo, tienen importantes efectos en los resultados. PLS-2 calibra para todos los componentes simultáneamente, es decir, los resultados de la descomposición espectral dan un conjunto de concentraciones y otro de vectores propios por calibración. Por lo tanto, los vectores calculados no están optimizados para cada componente individual sino para el conjunto de espectros. Esto puede suponer la pérdida de la exactitud en las predicciones de las concentraciones constituyentes, especialmente para medidas de muestras complejas.

En PLS-1 se calcula un conjunto separado de puntuaciones y vectores para cada componente de interés. En este caso, los conjuntos separados de vectores propios y concentraciones están calculados para cada componente y, por lo tanto, deberían ofrecer predicciones más exactas que PLS-2 [36].

La ventaja de usar PLS frente al modelo clásico de mínimos cuadrados es que realiza la descomposición espectral y el ajuste de todos los espectros al tiempo en un único paso. Los vectores propios están directamente relacionados con los parámetros de interés y una vez obtenidos se pueden utilizar para estimar cualquier muestra. Este modelo presenta mayor robustez de calibración, soporta un rango de variabilidad mayor frente a muestras desconocidas. El principal inconveniente es que requiere de un alto número de muestras y observaciones para obtener una calibración precisa y que el tiempo de cálculo es mayor al realizar la descomposición de todos los espectros simultáneamente.

### 5.1.2 Técnicas de procesamiento para clasificación cualitativa

Se utilizan técnicas basadas en la segmentación y/o umbralización. La segmentación consiste en la separación de diferentes regiones dentro de una misma imagen, en nuestro caso carne y grasa. Existen una gran variedad de métodos para la segmentación de medidas hiperespectrales, en este trabajo se han implementado los siguientes.

#### *Algoritmo de segmentación Otsu*

El algoritmo Otsu permite realizar una segmentación por umbralización cuando hay una diferenciación entre los objetos que se quieren distinguir. Se basa en estadísticas sobre el histograma unidimensional, es decir, la distribución de los valores de las muestras. Este método selecciona el umbral óptimo maximizando la varianza entre las clases.

$$g(x, y) = \begin{cases} 1 & \Leftrightarrow f(x, y) > T \\ 0 & \Leftrightarrow f(x, y) \leq T \end{cases}$$

Este método tiene buena respuesta a la mayoría de las situaciones y no precisa de supervisión humana ya que, de forma automática gracias a un preprocesamiento de la imagen e información relacionada, puede establecer el umbral de forma adaptativa.

El método funciona como sigue: si se tiene una imagen con diferentes valores, formada por  $N$  píxeles y  $L$  posibles valores diferentes, la intensidad media total de la imagen se calcula a partir de la media de cada una de las clases que luego Otsu utiliza para examinar la varianza se define como:

$$\begin{aligned} \omega_1 \mu_1 + \omega_2 \mu_2 &= \mu_T \\ \omega_1 + \omega_2 &= 1 \end{aligned}$$

Otsu utilizando un análisis discriminante, define la varianza entre clases de una imagen umbralizada como:

$$\sigma_B^2 = \omega_1 (\mu_1 - \mu_T)^2 + \omega_2 (\mu_2 - \mu_T)^2$$

Y lo que se pretende conseguir es localizar el valor donde sea máxima dicha varianza:

$$\begin{aligned} t^* &= \max_t \{\sigma_B^2(t)\} \\ 1 &\leq t \leq L \end{aligned}$$

Una vez explicado como funciona el algoritmo Otsu, en este trabajo se aplica sobre una matriz que contiene datos significativos del espectro para pixel de la muestra, adquiridos mediante diferentes métodos que se describen a continuación.

- MIP, Proyección de Máxima intensidad (*Maximum Intensity Projection*): selecciona de la matriz tridimensional con toda la información espectral  $R(x, y, \lambda)$  el valor máximo de la reflectancia del material una vez se ha corregido espectralmente.
- invMIP, Proyección de Mínima intensidad: selecciona de la matriz tridimensional con toda la información espectral  $R(x, y, \lambda)$  el valor mínimo de la reflectancia del material una vez se ha corregido espectralmente.
- Promedio: selecciona de la matriz tridimensional con la información espectral  $R(x, y, \lambda)$  el valor promedio de una dimensión, en este caso la dimensión espectral.

### ***Método de segmentación utilizando la distancia euclídea***

Método de segmentación mediante el cálculo de la distancia euclídea entre espectros. Se comparan los espectros de cada punto de la muestra con uno o varios espectros de referencia para obtener así una matriz de escalares que indican el grado de similitud respecto a la variable de referencia. En este proyecto las referencias son 3: muestra de grasa pura, muestra de carne novilla pura y muestra de carne de cerdo pura, es decir, concentración al 100%. La diferencia euclídea se calcula a través de la siguiente expresión.

$$D(\lambda) = \frac{1}{2} \|R(\lambda) - r(\lambda)\|^2$$

donde  $D(\lambda)$  es el resultado de la distancia euclídea,  $R(\lambda)$  es la reflectancia difusa de la muestra por clasificar y  $r(\lambda)$  es el espectro de referencia [33].

### *Método de segmentación mediante Mapeo del Ángulo Espectral (SAM)*

El procedimiento del método SAM (*Spectral Angle Mapper*) es muy similar al anterior. Se realiza también la comparación entre los espectros de cada punto de la muestra con los de referencia de carne y grasa, pero en vez de obtener la distancia euclídea entre ambos se calcula el ángulo espectral, es decir el ángulo que forman los vectores con la información del espectro.

$$\vec{e}_i(\lambda) \cdot \vec{e}_{ref}(\lambda) = |\vec{e}_i(\lambda)| \cdot |\vec{e}_{ref}(\lambda)| \cdot \cos\theta$$
$$\theta = \cos^{-1} \frac{\vec{e}_i(\lambda) \cdot \vec{e}_{ref}(\lambda)}{|\vec{e}_i(\lambda)| \cdot |\vec{e}_{ref}(\lambda)|}$$

donde  $\vec{e}_i(\lambda)$  es el espectro a evaluar,  $\vec{e}_{ref}(\lambda)$  es el espectro de referencia y  $\theta$  es el ángulo que los separa en el espacio N-dimensional correspondiente a las N longitudes de onda ( $\lambda$ ) en las que se mide la reflectancia del material. Como se puede comprobar, el cálculo del ángulo se realiza a partir de la expresión del producto escalar entre ambos vectores, como se expresa en la ecuación anterior.



## Parte IV: Resultados



# Capítulo 6

## *Clasificación cuantitativa*

Este capítulo tiene como objetivo analizar los resultados obtenidos utilizando las diferentes técnicas de procesamiento para la clasificación cuantitativa expuestas en el capítulo anterior, con las que se pretende estimar el porcentaje de grasa de cualquier hamburguesa.

Los dos métodos utilizados son, el método clásico de mínimos cuadrados (CLS, *Classical Least Squares*) y el método mínimos cuadrados parciales (PLS, *Partial Least Squares*), que han sido explicados en la teoría del apartado anterior.

### 6.1. Método Clásico de Mínimos Cuadrados

Este método de cuantificación a través del coeficiente de correlación de una serie de medidas y su aproximación lineal está relacionado con la longitud de onda a la que se observan los valores de dichas las medidas. Por lo tanto, la bondad del coeficiente de correlación depende de la longitud de onda.

#### Obtención espectros promedio

Para poder obtener una aproximación por mínimos cuadrados, el primer paso, tanto para carne de cerdo como para carne de novilla, es leer la información espectral en cada punto, recorriendo pixel por pixel cada una de las 8 mezclas con diferentes porcentajes de grasa, calculando su espectro promedio. El promedio se calcula tanto de la región VIS-NIR, como de la región SWIR, para la carne de cerdo y para la carne de novilla, plasmando de ese modo los resultados en dos gráficas diferentes. Por un lado, la carne de cerdo, en ambas regiones, VIS-NIR y SWIR, y, por otro lado, la carne de novilla.



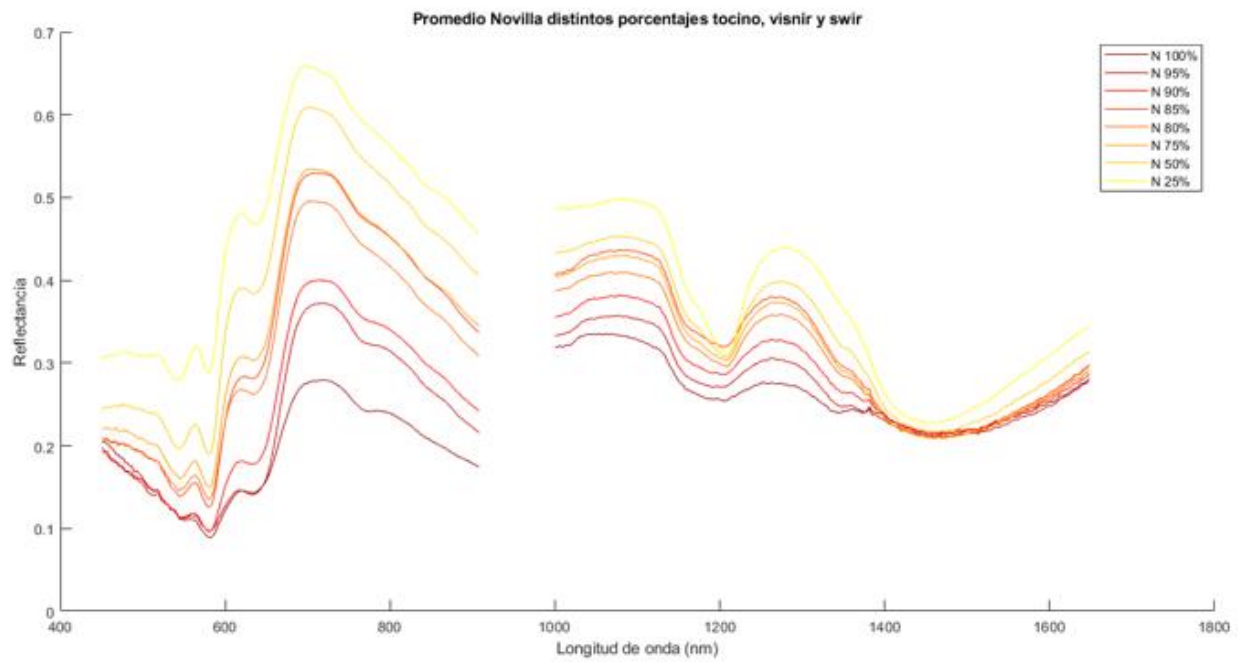


Figura 17: Espectros promedio de las muestras de carne de novilla con distintos porcentajes de grasa.

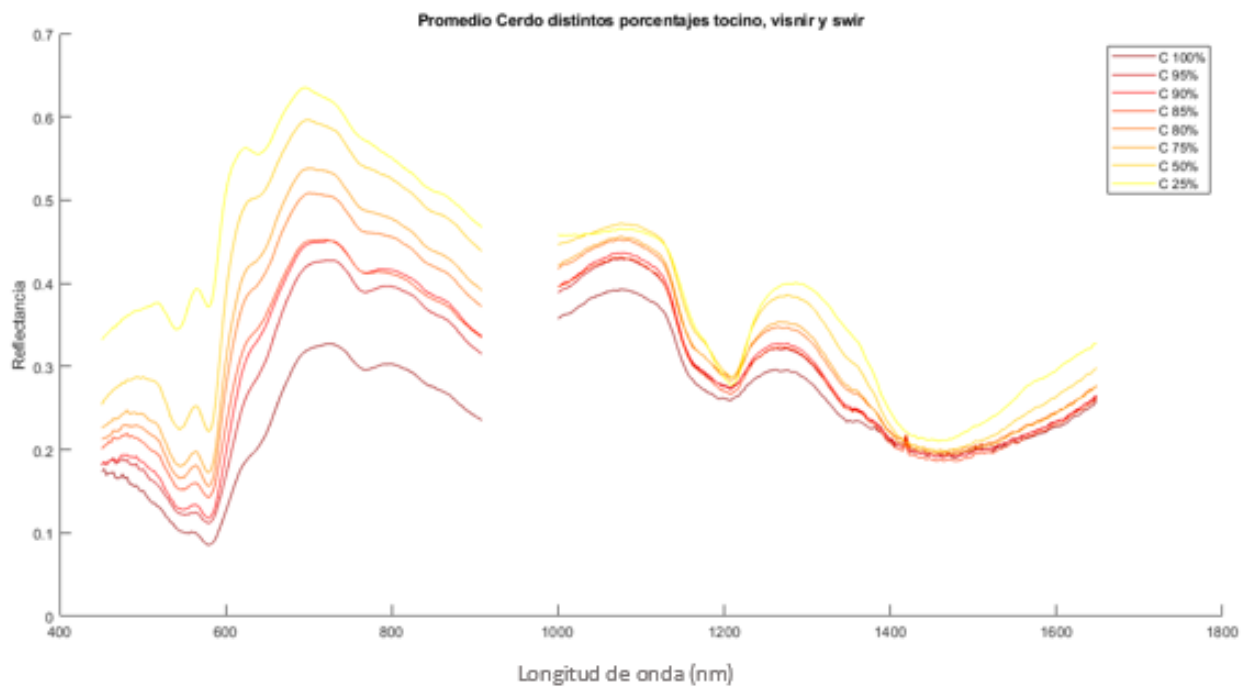


Figura 18: Espectros promedio de las muestras de carne de cerdo con distintos porcentajes de grasa

## Diferencia espectral carne de novilla vs carne de cerdo

Para poder apreciar mejor las diferencias entre los dos tipos de carne, la de novilla y la de cerdo, se representan los espectros promedios de las muestras al 100% de carne, es decir, sin grasa añadida. Como se puede ver en la Figura 19, se obtienen los resultados esperados. La grasa de cerdo muestra valores más altos de reflectancia en ambos sistemas, tanto VIS-NIR como en SWIR, porque la carne de cerdo de por sí contiene más grasa que la de novilla.

En la zona SWIR, en torno 1400 nm se produce un cruce entre ambas gráficas. Esto puede darse debido a que la carne de cerdo tiene más grasa que la de novilla, y coincide con las longitudes de onda donde se sitúa el pico de absorción principal teórico del agua y cercano al de los lípidos. Por tanto, esto produce una mayor absorción en la carne de cerdo y por tanto los valores de reflectancia son más bajos en ese pico de absorción.



Figura 19: Espectros promedio de las muestras de novilla 100% de carne y cerdo 100% de carne

## Localización de las zonas con mayor linealidad del espectro

Buscando cuál es la mejor forma de representar los datos de reflectancia de las muestras, se comparan las gráficas de los espectros promedio de las reflectancias ya calibradas que mide directamente el sistema, con la representación de la reflectancia, aplicando el logaritmo neperiano o el logaritmo decimal, para ver si así mejora la linealidad y por tanto el coeficiente de correlación para cada longitud de onda entre los datos medidos de los espectros promedio y la aproximación lineal por mínimos cuadrados.

El objetivo de aplicar logaritmos a los valores de reflectancia es relacionar directamente la información espectral con la concentración de cromóforos (grasa, colágeno, hemoglobina...) a través de la nueva variable  $X(\lambda)$  que se observa en las siguientes fórmulas:

$$R(\lambda) = A\lambda^{-b} \exp \left[ -\rho C_{HbT} \left( f_{O_2} \varepsilon_{HbO_2}(\lambda) + (1 - f_{O_2}) \varepsilon_{Hb}(\lambda) \right) \right] \quad (19)$$

$$\begin{aligned} X(\lambda) &= \ln(R(\lambda)) \\ &= \ln(A) - b \ln(\lambda) - \rho C_{HbT} \left( f_{O_2} \varepsilon_{HbO_2}(\lambda) + (1 - f_{O_2}) \varepsilon_{Hb}(\lambda) \right) \\ &= S_1 \sigma_1(\lambda) + S_2 \sigma_2(\lambda) + \dots + S_n \sigma_n(\lambda) \end{aligned} \quad (20)$$

Donde  $A$  es la amplitud de scattering,  $b$  la potencia de scattering,  $\rho$  el recorrido de la luz,  $C_{HbT}$  la concentración de hemoglobina,  $f_{O_2}$  la fracción de hemoglobina oxigenada,  $\varepsilon_{HbO_2}(\lambda)$  y  $\varepsilon_{Hb}(\lambda)$  son el coeficiente de extinción molar de la hemoglobina oxigenada y desoxigenada respectivamente.  $X(\lambda)$  es el logaritmo de la reflectancia  $R(\lambda)$  donde las concentraciones del espectro son combinaciones lineales.  $S_1, S_2$  y  $S_3$  son los pesos lineales que modulan la contribución de cada componente espectral, donde cada comportamiento espectral es representado por las funciones  $\sigma_1(\lambda)$ ,  $\sigma_2(\lambda)$  y  $\sigma_3(\lambda)$  directamente relacionadas con la absorción. [37]

En el capítulo 2, se ha explicado dónde se sitúan los picos de absorción más significativos de los principales cromóforos presentes en la carne. Para comprobar si las zonas de mayor linealidad de los espectros promedio medidos para los distintos porcentajes de grasa coinciden con los picos de absorción teóricos, o si por el contrario las zonas de mayor linealidad se encuentran en otras longitudes de onda, en la Figura 20 se representa el coeficiente de relación al aplicar la aproximación lineal de mínimos cuadrados en cada punto del espectro. Se realiza tanto para la representación de la reflectancia en las tres formas mencionadas previamente, reflectancia, logaritmo neperiano de la reflectancia y logaritmo decimal de la reflectancia

## Comparación linealidad carne de novilla

### ○ Reflectancia $R(\lambda)$

En la Figura 20, la gráfica superior representa los espectros promedios de la carne de novilla donde la leyenda indica los porcentajes de grasa. La gráfica inferior muestra el coeficiente de correlación conseguido en el modelo de calibración basado en CLS para cada longitud de onda. El coeficiente de correlación se obtiene a partir de los 8 espectros promedio en cada longitud de onda concreta. Se realiza una aproximación lineal a esos 8 valores y se compara el grado de linealidad del ajuste por mínimos cuadrados clásico.

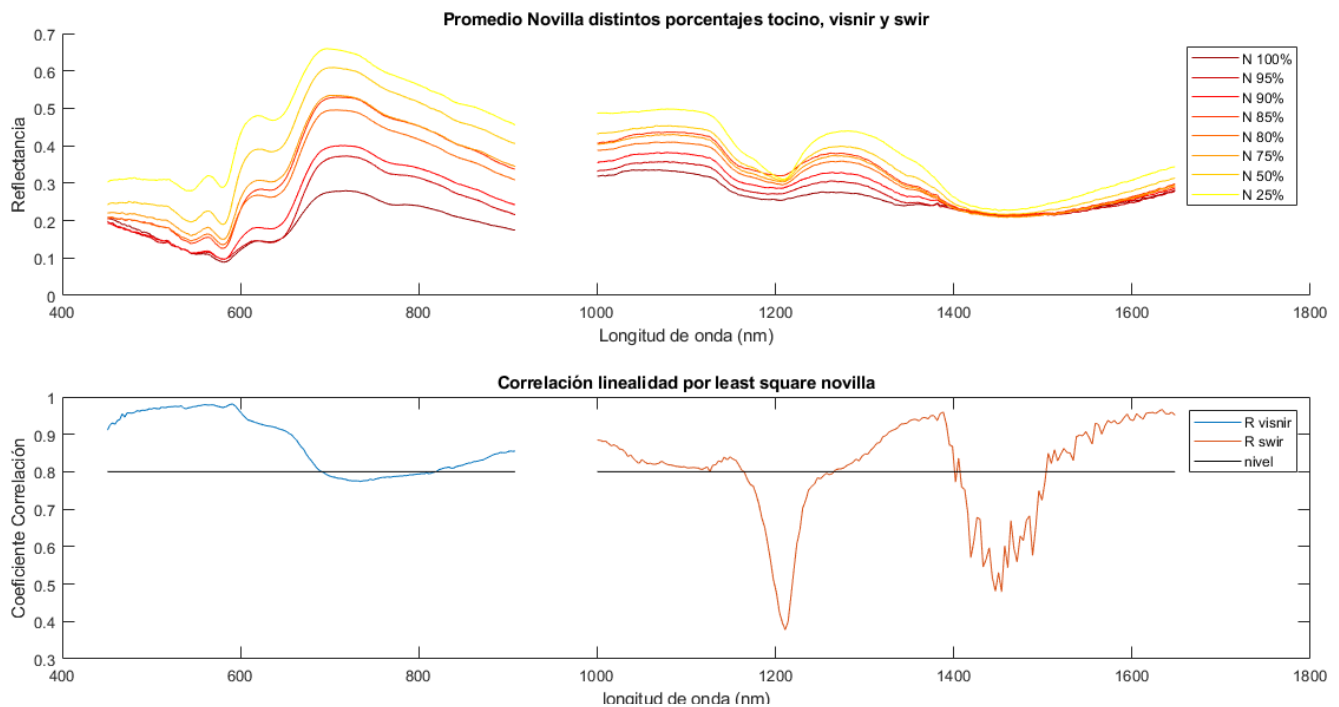


Figura 20: Espectros promedio y coeficiente de correlación de las muestras de carne de novilla con distintos porcentajes de grasa

### ○ Reflectancia $\ln(R(\lambda))$ y $\log(R(\lambda))$

Se quiere comprobar si el coeficiente de correlación mejora si los valores de la reflectancia se expresan aplicando el logaritmo neperiano y decimal. Al calcular el coeficiente de correlación del logaritmo neperiano de la reflectancia y compararlo con el logaritmo decimal de la reflectancia se observa que coinciden, son exactamente iguales. A través del coeficiente de correlación se busca comprobar si expresar la reflectancia como  $\ln(R(\lambda))$  mejora la linealidad por estar relacionada directamente con la concentración de cromóforos.

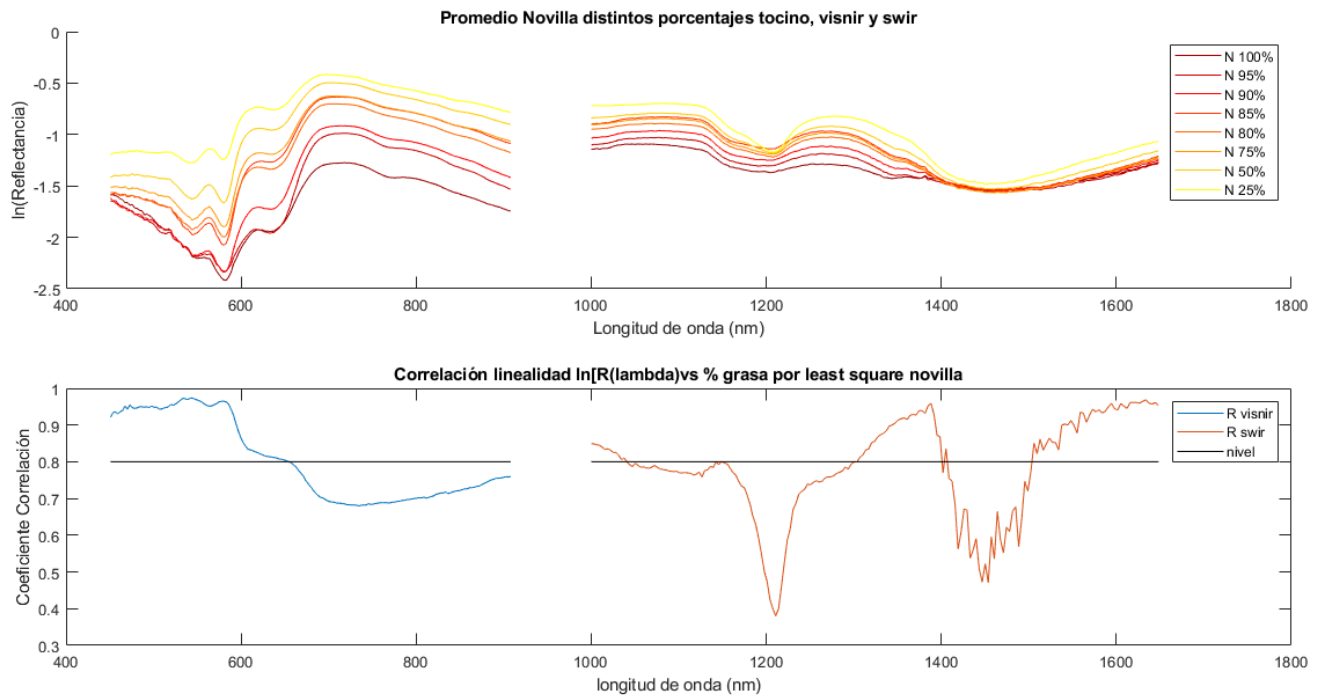


Figura 21: Espectros promedio y coeficiente de correlación de las muestras de carne de novilla con distintos porcentajes de grasa (logaritmo)

Como se puede observar al aplicar logaritmos el coeficiente de relación no mejora sustancialmente para la carne de novilla obteniéndose valores más bajos de  $R^2$  en todas las longitudes de onda.

### Comparación linealidad carne de cerdo

#### ○ Reflectancia $R(\lambda)$

Se representan los espectros promedio de las muestras de carne de cerdo en función de los valores de reflectancia directamente medidos y calibrados por los sistemas. Al calcular el coeficiente de correlación para esos valores de reflectancia para cada longitud de onda se observa que sigue un comportamiento parecido al coeficiente de correlación de la carne de novilla, desplomándose también, en la región SWIR coincidiendo con los picos de absorción de 1200 y 1400 nm.

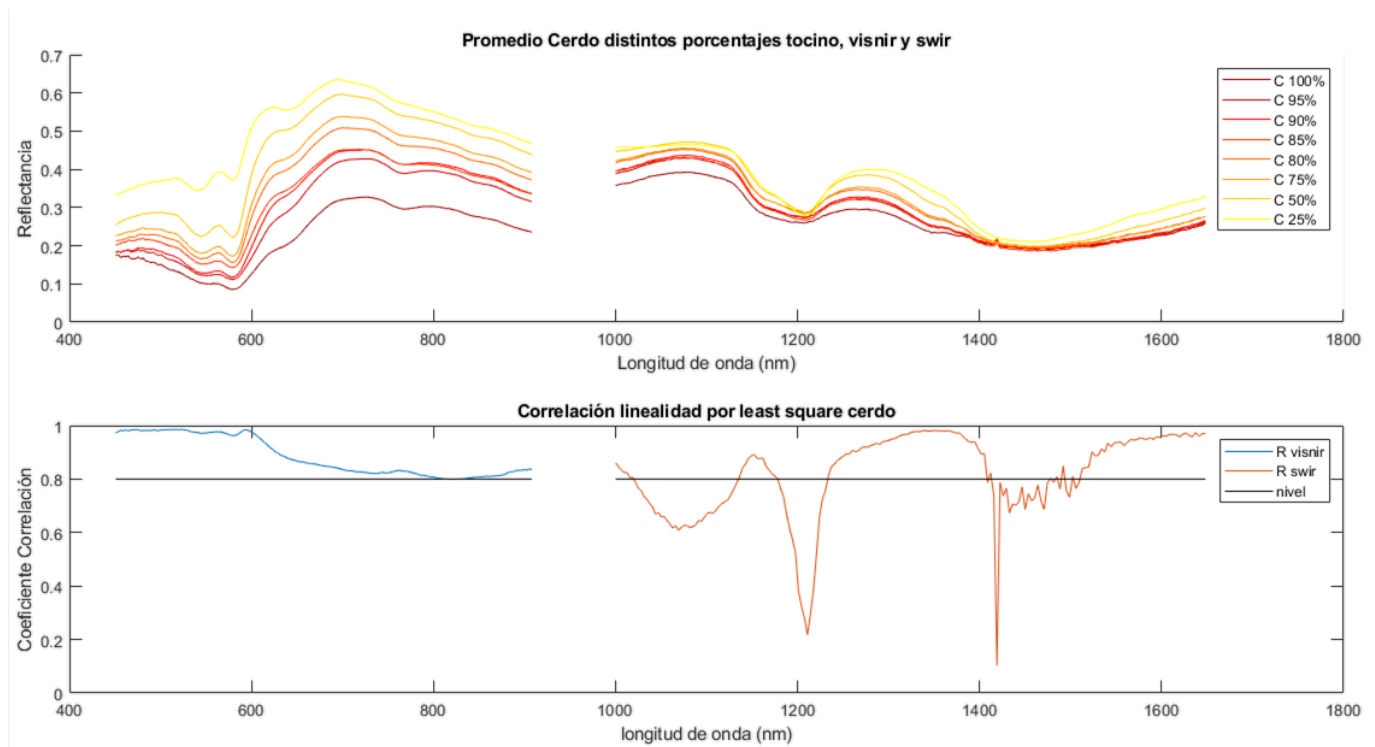


Figura 22: Espectros promedio y coeficiente de correlación de las muestras de carne de cerdo con distintos porcentajes de grasa

○ Reflectancia  $\ln(R(\lambda))$  y  $\log(R(\lambda))$

Como ocurre en el caso de la carne de novilla aplicar logaritmos a  $R(\lambda)$  tampoco mejora la linealidad de los espectros promedio de la carne de cerdo, obteniéndose valores del coeficiente de correlación más bajos en la región VIS-NIR, y aunque en la región SWIR la diferencia no es tan evidente, los valores también son ligeramente inferiores.

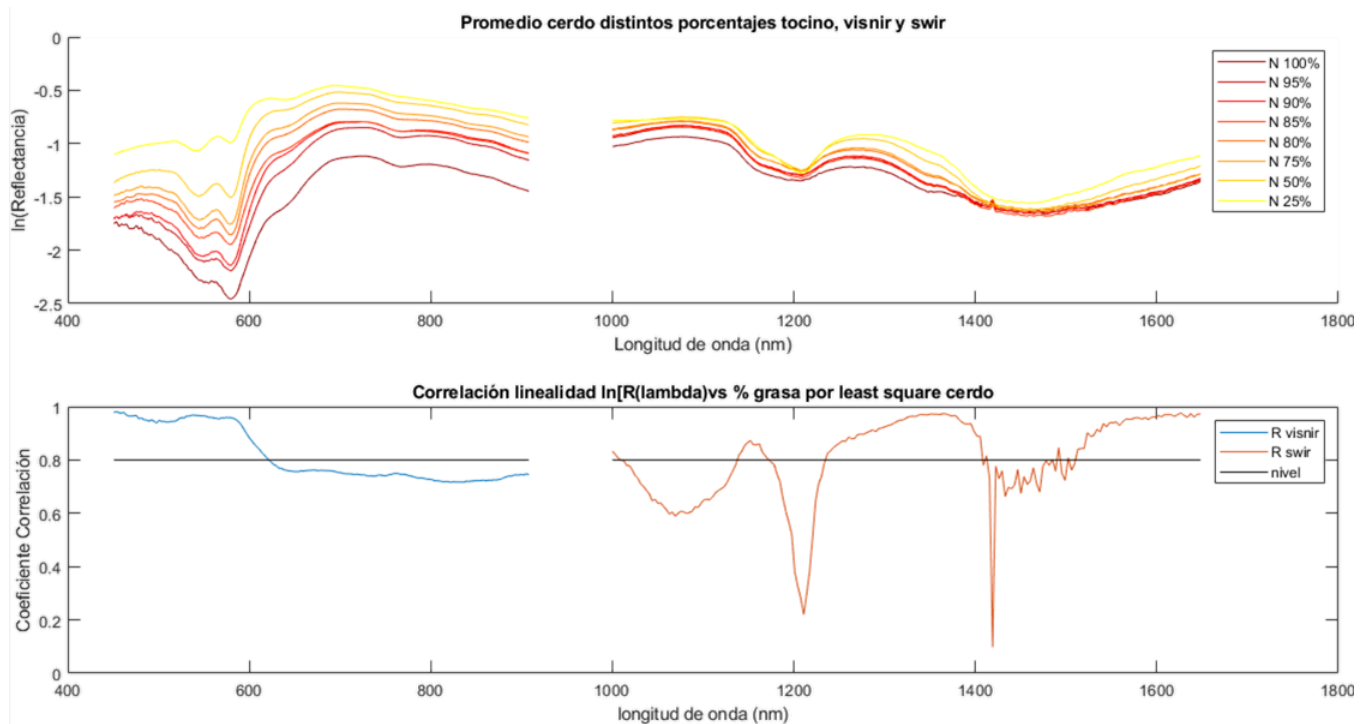


Figura 23: Espectros promedio y coeficiente de correlación de las muestras de carne de cerdo con distintos porcentajes de grasa (logaritmo)

### Comparación general linealidad novilla y cerdo

Para ambos tipos de carne se aprecia que en la región VIS-NIR hay una correlación alta cercana al valor máximo de  $R^2=1$  en los picos de absorción medidos en las longitudes de onda de 544nm y 580nm que se corresponden con los picos de absorción teóricos de la hemoglobina oxigenada. En cambio, en la región SWIR, en los picos de absorción característicos del agua y de los lípidos, que se sitúan en torno a 1200 y 1400nm, se pierde esa linealidad, se puede apreciar en todas las gráficas, observando la zona donde el coeficiente de correlación cae bruscamente. Una de las hipótesis para este fenómeno es que la absorción del agua es tan significativa que enmascara el resto de las características espectrales de la composición de la muestra. Pudiendo concluir que, en la región SWIR, la linealidad no se corresponde con las zonas del espectro donde se sitúan los picos de absorción teóricos. Esa pérdida de la linealidad se puede deber a que las muestras con mayor porcentaje de grasa tienen un coeficiente de absorción mayor y se entremezclan con los valores de reflectancia de las muestras con proporción de carne más alta. Otra de las razones por las que se puede perder la linealidad, es que la grasa, fuera de sus correspondientes picos de absorción, refleja más y puede producir más fácilmente esparcimiento (*scattering*), afectando así a la medida realizada.

## Comparación de la linealidad en las longitudes de onda del espectro con mayor coeficiente de correlación

Una vez identificadas las zonas del espectro con una linealidad superior a un umbral establecido ( $R^2 = 0.8$ ), se procede a observar la aproximación lineal por mínimos cuadrados en longitudes de onda concretas, cuantificando así si la correlación es más alta en la carne de cerdo o novilla, qué método de representación de la reflectancia es mejor y el comportamiento cuando omitimos en la estimación un punto que corresponde a una muestra que podría no tener el porcentaje de grasa previsto por el posible mal error en la preparación de las muestras.

### ○ Hemoglobina oxigenada

Para analizar más detalladamente los resultados del ajuste, se va a estudiar el caso de la hemoglobina oxigenada alrededor de su pico de absorción teórico a 576nm.

#### - Reflectancia $R(\lambda)$

En la Figura 24 se compara el ajuste lineal de las medidas hechas sobre la carne de cerdo con las de la novilla, para el caso de la reflectancia expresada como  $R(\lambda)$ . Aunque ambas medidas se parecen mucho al ajuste, se aprecia una desviación significativa en la muestra de 50% de carne, siendo mayor en el caso del cerdo. El coeficiente de correlación es mayor en el caso de la novilla.

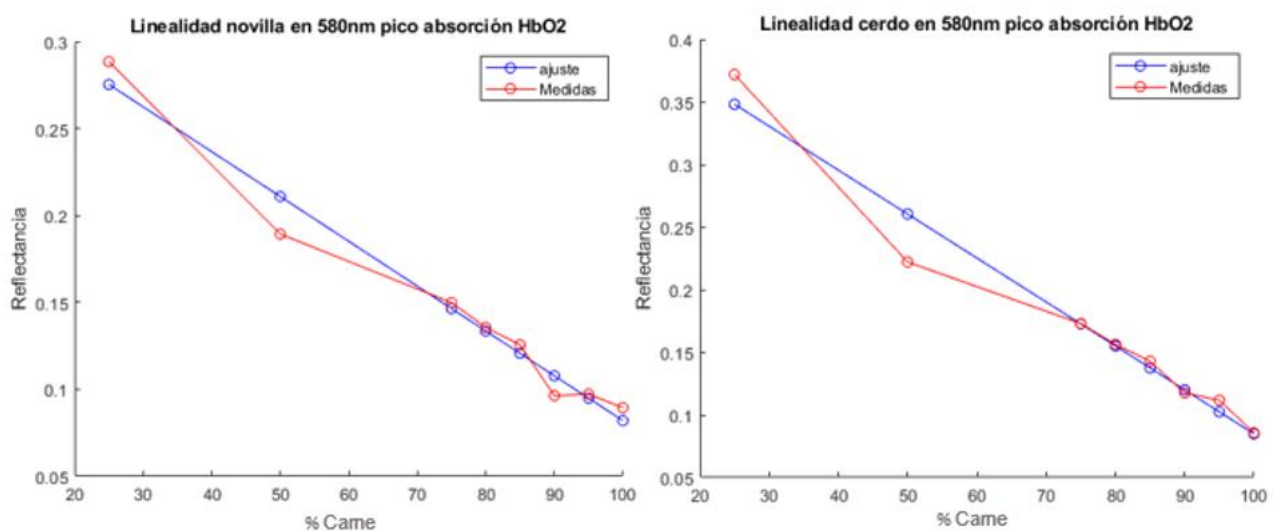
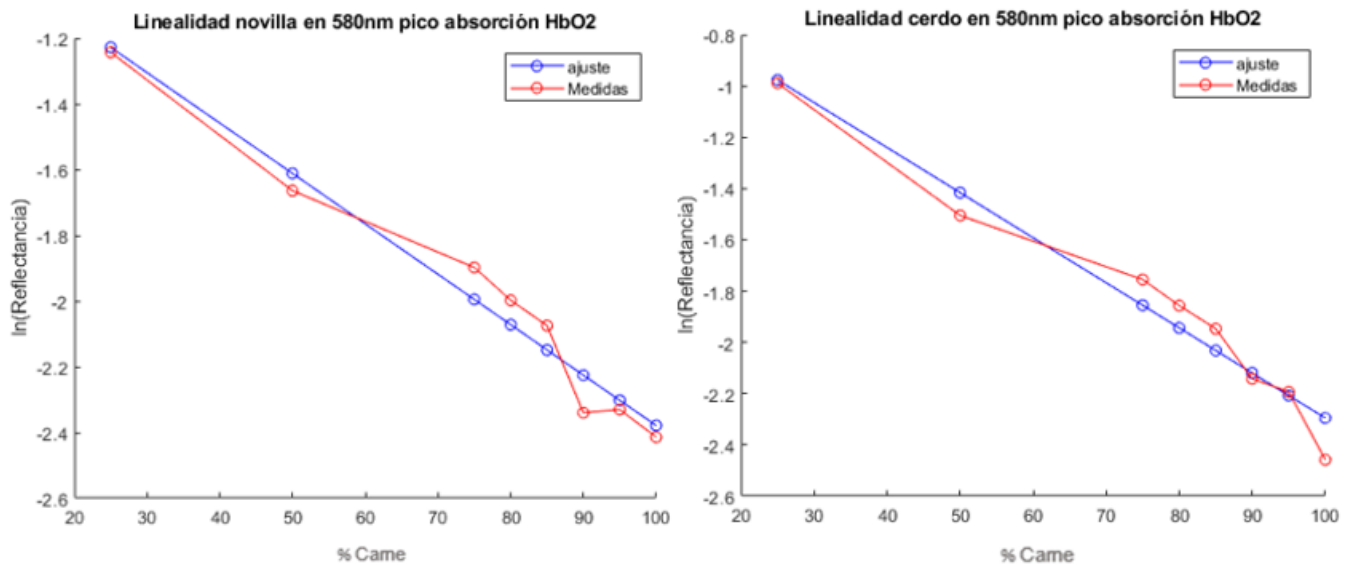


Figura 24: Comparativa de linealidad a 580nm pico absorción HbO2 en las carnes de novilla y cerdo



### - Reflectancia $\ln(R(\lambda))$

Se quiere comprobar cuanto varía el coeficiente de correlación y el ajuste lineal, al aplicar la reflectancia en función del logaritmo neperiano. En ambos casos, la aproximación sigue siendo buena, pero no se aproxima tanto al valor real de las muestras, como en el caso de la reflectancia  $R(\lambda)$ .



**Figura 25: Comparativa de linealidad a 580nm pico absorción HbO2 en logaritmo de las carnes de novilla y cerdo**

Como ya se ha visto en las Figuras 21 y 23 anteriores, el ajuste por mínimos cuadrados aplicando el logaritmo no mejora los datos del coeficiente de correlación. Esto queda constatado tanto en las figuras anteriores como en la siguiente tabla.

Se puede concluir que existe una mayor linealidad en la carne novilla que en la de cerdo, posiblemente porque la carne de novilla de por sí tiene mayor cantidad de carne, y la del cerdo tiene más grasa, por lo que su comportamiento es menos predecible debido al esparcimiento (*scattering*) que produce.

$R^2$	$R(\lambda)$	$\ln(R(\lambda))$
Novilla	0.9718	0.9649
Cerdo	0.9627	0.9585

**Tabla 6: Comparativa coeficiente de correlación sobre novilla y cerdo a 580nm**

### ○ Novilla: Lípidos 1040nm

Se va a estudiar el comportamiento lineal en el caso de las muestras de carne de novilla en la longitud de onda de 1040 nm relacionada con el pico de absorción de los lípidos. Se va a analizar en 1040 nm, porque el pico característico de absorción de lípidos a 1200 nm se sitúa dentro de una de las zonas donde el coeficiente de correlación empeora.

Como también se puede apreciar en las Figuras anteriores 22 y 23, en el caso de la carne de cerdo, el coeficiente de correlación está bastante por debajo del nivel ( $R^2 = 0.8$ ), por lo que sólo se van a discutir los resultados para la carne de novilla.

Si observamos la Figura 26 y lo que sucede en el caso de la hemoglobina oxigenada para la carne de novilla, la muestra con un 90% de carne se desvía del rango de valores esperados si la mezcla contuviese esas proporciones, por lo que se compara el método de ajuste de mínimos cuadrados teniendo en cuenta o no ese punto.

#### - Reflectancia $R(\lambda)$

Comparando la siguiente figura, se puede observar cómo mejora la aproximación lineal, por tanto, el coeficiente de correlación, cuando no se tiene en cuenta la muestra del 90%.

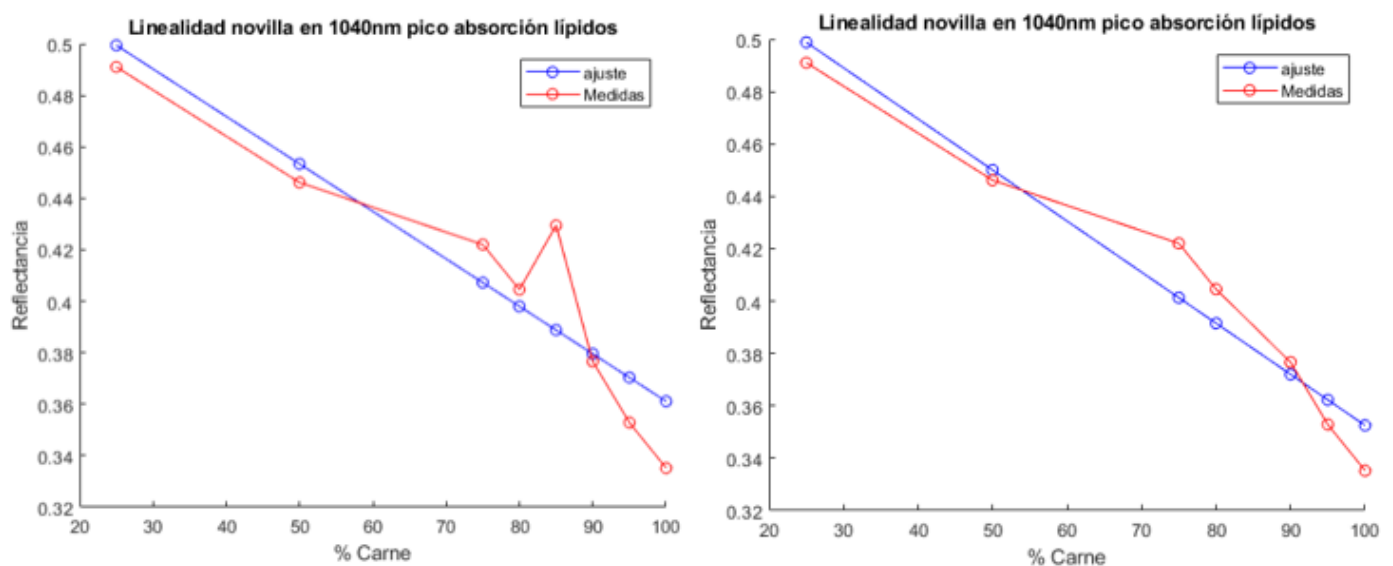


Figura 26:Comparativa de linealidad a 1040nm pico absorción de lípidos en de las carnes de novilla con y sin muestra al 90%

#### - Reflectancia $\ln(R(\lambda))$

Comparando con los resultados de la reflectancia expresada como  $R(\lambda)$  (Figura 26), al aplicar logaritmos no se mejora el coeficiente de correlación, y sigue siendo mejor omitir en el ajuste la muestra del 90% de carne.

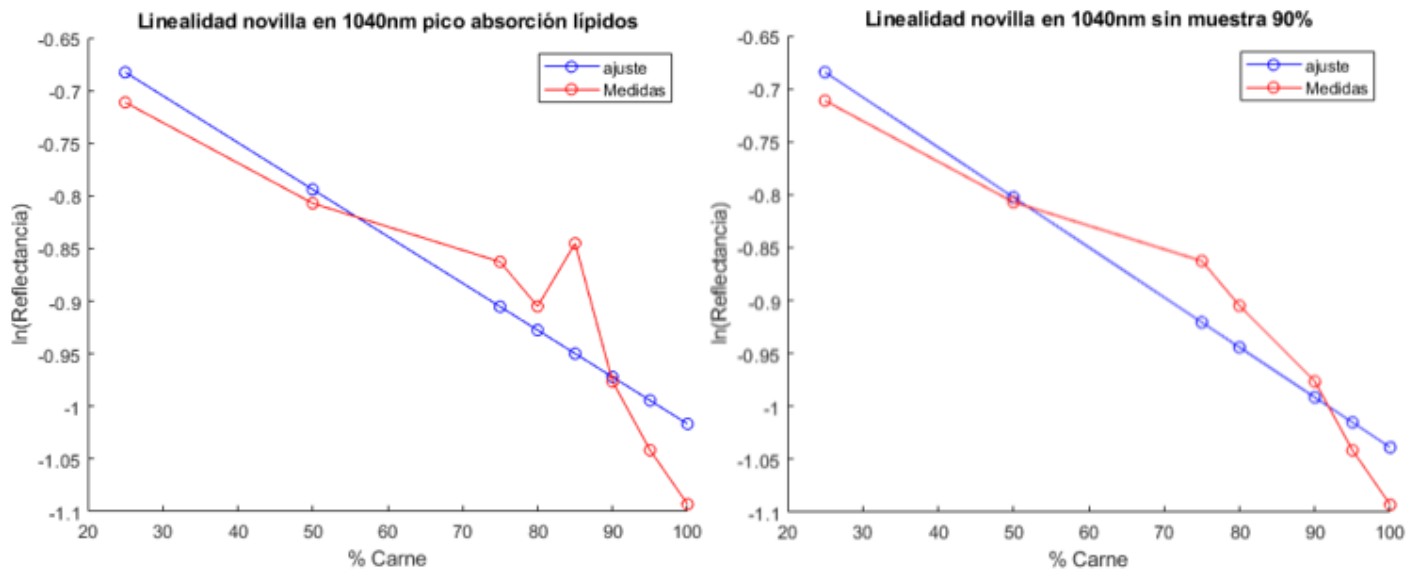


Figura 27: Comparativa de linealidad a 1040nm pico absorción de lípidos en logaritmo de la carne de novilla con y sin muestra al 90%

En la siguiente tabla se recogen los coeficientes de correlación de todos los casos que se han comparado anteriormente, para conocer numéricamente el grado de aproximación del ajuste lineal.

$R^2$	$R(\lambda)$	$\ln(R(\lambda))$
Novilla	0.8350	0.7997
Novilla sin 90%	0.9390	0.9115

Tabla 7: Comparativa correlación en carne de novilla a 1040nm

Como discusión, la desviación presente en el punto del 90% de carne, que provoca un coeficiente de correlación más bajo, puede ser debida a que a la hora de preparar las muestras, la porción depositada en la cubeta que posteriormente se analizó no contenía realmente esas proporciones. La mezcla contiene un porcentaje de grasa diferente, que da lugar a esos valores de reflectancia. Todo esto se debe al método de realización, en el que la mezcla fue preparada y compactada manualmente, lo que conlleva a una probabilidad más alta de error.

#### - Punto de mayor linealidad

Se ha identificado la longitud de onda donde el  $R^2$  es máximo, es decir donde hay mayor linealidad y por tanto existirá menos error al estimar el porcentaje de grasa en esas longitudes de onda.

El valor de  $R^2$  máximo se ha obtenido recorriendo las regiones VIS-NIR y SWIR en las Figuras 20 y 21, tanto para la reflectancia  $R(\lambda)$  como  $\ln(R(\lambda))$ . Se encuentran representados en la Tabla

8 los valores máximos de los coeficientes de correlación y los diagramas de linealidad en la Figura 28.

VIS-NIR	$R^2$	$R(\lambda)$	$\ln(R(\lambda))$
		0.9806	0.9735
	Longitud de onda (nm)	591	544
SWIR	$R^2$	$R(\lambda)$	$\ln(R(\lambda))$
		0.9667	0.9618
	Longitud de onda (nm)	1633	1615

Tabla 8. Coeficientes de correlación máximos y sus respectivas longitudes de onda. Carne de Novilla

Como se puede observar en el caso de  $\ln(R(\lambda))$  en la región VIS-NIR la longitud de onda a la que el coeficiente de correlación es máximo coincide con el pico de absorción a 544nm de la hemoglobina oxigenada.

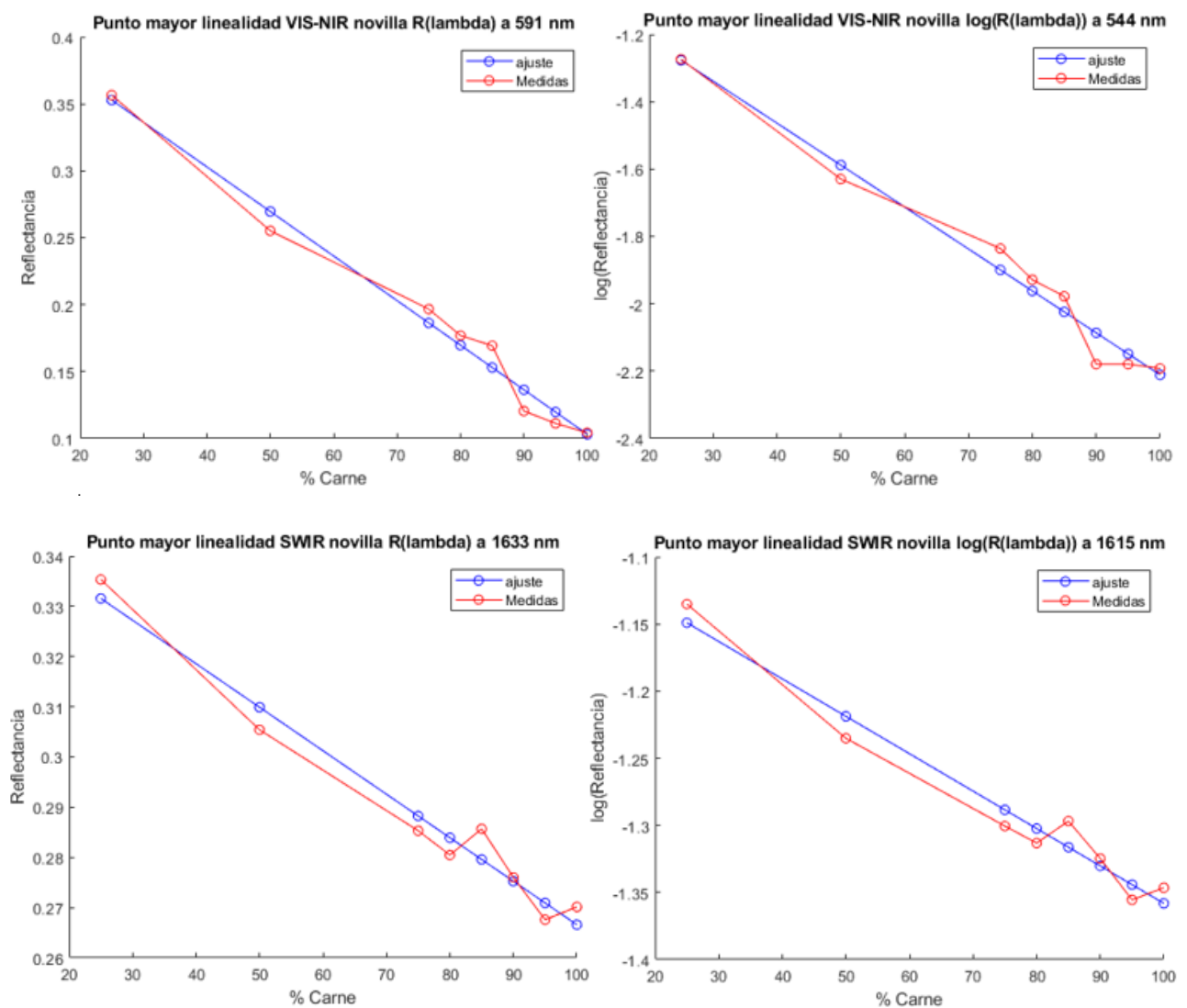


Figura 28. Gráficas de linealidad de las longitudes de onda donde los coeficientes son máximos en la región VIS-NIR como SWIR carne de novilla

### ○ Cerdo: Pico absorción grasa de cerdo 762 nm

Se quiere comprobar si la desviación observada para la muestra de cerdo con porcentaje de 85% de carne afecta al coeficiente de correlación final. Para ello se va a analizar el comportamiento que tiene a 762 nm, longitud de onda, que corresponde con un pico de absorción de grasa de cerdo.

#### - Reflectancia $R(\lambda)$

Comparando la Figura 29, se observa que el ajuste prácticamente no varía al omitir la muestra. De hecho, se obtienen mejores resultados si no se omite.

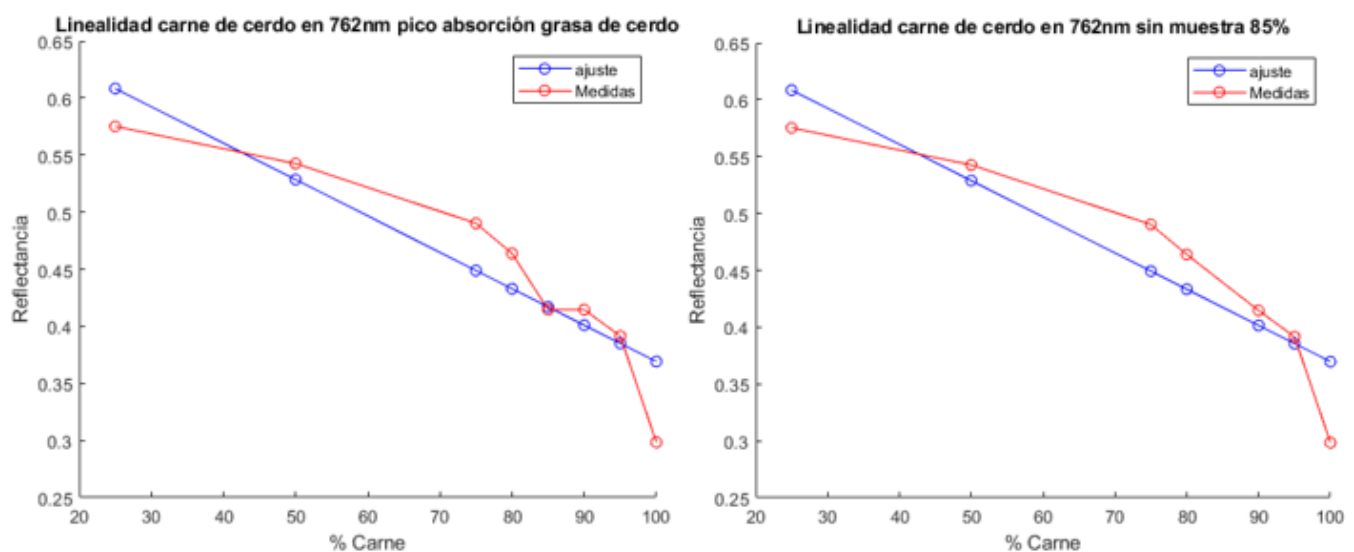


Figura 29: Comparativa de linealidad a 762nm pico absorción de la grasa en la carne de cerdo con y sin muestra al 85%

#### - Reflectancia $\ln(R(\lambda))$

Al expresar la reflectancia con el logaritmo neperiano, los resultados obtenidos empeoran, aunque siguen el mismo comportamiento que en el caso de  $R(\lambda)$ , es decir, es mejor no omitir la muestra.

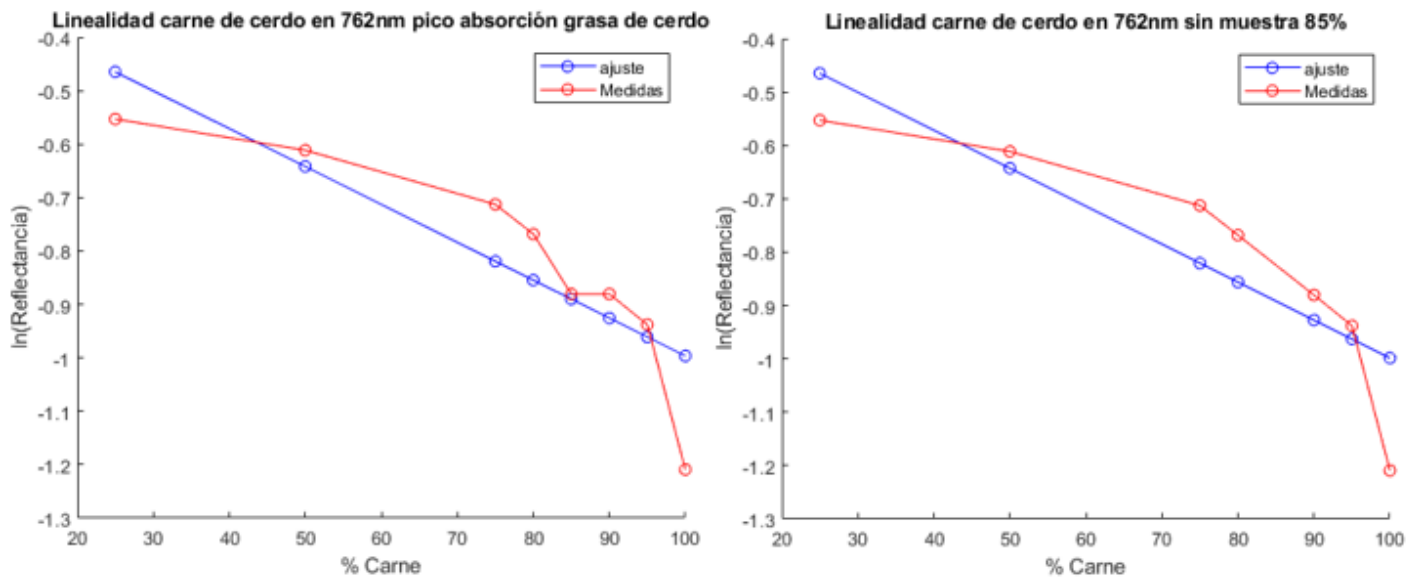


Figura 30: Comparativa de linealidad a 762nm pico absorción de la grasa en la carne de cerdo en logaritmo con y sin muestra al 85%

Estas conclusiones quedan reforzadas con los coeficientes de correlación que se muestran en la siguiente tabla.

$R^2$	$R(\lambda)$	$\ln(R(\lambda))$
Cerdo	0.8328	0.7499
Cerdo sin 85%	0.8287	0.7466

Tabla 9: Comparativa correlación en carne de cerdo a 762nm

#### - Punto de mayor linealidad

Como en el caso de la novilla se han identificado las longitudes de onda donde el  $R^2$  es máximo, para poder realizar la clasificación cuantitativa de la grasa en los puntos del espectro más lineales de las regiones VIS-NIR y SWIR. El valor de  $R^2$  máximo se ha obtenido recorriendo las regiones VIS-NIR y SWIR en las Figuras 22 y 23, tanto para la reflectancia  $R(\lambda)$  como  $\ln(R(\lambda))$ . Se encuentran representados en la Tabla 10 los valores máximos de los coeficientes de correlación y los diagramas de linealidad en la Figura 31.

VIS-NIR	$R^2$	$R(\lambda)$	$\ln(R(\lambda))$
	Longitud de onda (nm)	522	453
SWIR	$R^2$	$R(\lambda)$	$\ln(R(\lambda))$
	Longitud de onda (nm)	1340	1626

Tabla 10 Coeficientes de correlación máximos y sus respectivas longitudes de onda. Carne de Novilla

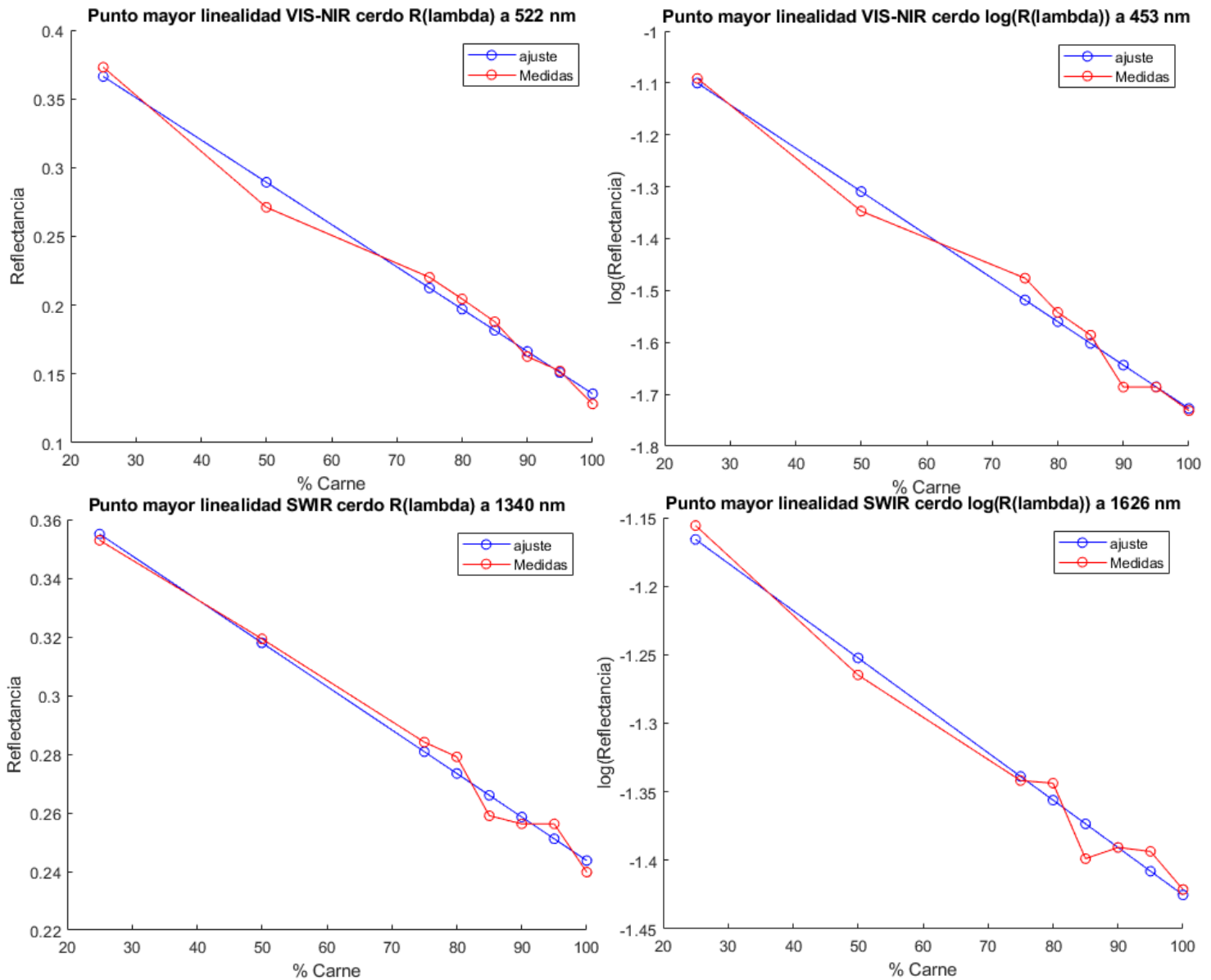
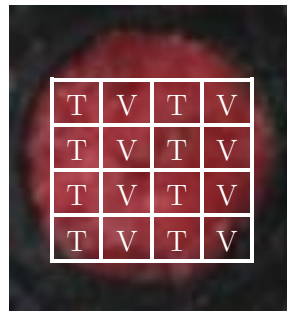


Figura 31 Gráficas de linealidad de las longitudes de onda donde los coeficientes son máximos en la región VIS-NIR como SWIR carne de cerdo

## 6.2. Método de Vectores Propios. PLS

PLS es un método de clasificación por vectores propios tal y como se explica en el capítulo 5 de metodología. Se basa en la estadística de todas las observaciones que sirven como entramiento. La función PLS determina cuales de esas observaciones, en este caso vectores que contienen la información espectral de cada pixel de la muestra usado para en la parte de entrenamiento, son más relevantes por sus parámetros estadísticos, por ejemplo, sus varianzas significativas. A partir de esos vectores, y sus correspondientes coeficientes, se puede clasificar cualquier muestra en función de su porcentaje de grasa.

Para entrenar el sistema se utiliza la mitad de los espectros de cada muestra y para validarlo la otra mitad. La selección se lleva a cabo como se muestra en la Figura 32, cogiendo pixeles de forma alterna para tener observaciones a la hora de entrenar el sistema que abarquen la mayor región de la muestra posible, cubriendo todas las zonas de forma que de existir alguna heterogeneidad significativa se tenga en cuenta en la fase de entrenamiento.



**Figura 32: Distribución de las zonas de entrenamiento y validación en la carne**

El proceso de clasificación mediante PLS de hamburguesas tanto de carne de cerdo como de novilla para las medidas realizadas con los sistemas que capturan el espectro tanto en la región VIS-NIR como SWIR sigue los mismos pasos, descritos a continuación.

La matriz de observaciones almacena por filas la información espectral de todos los pixeles de entrenamiento seleccionados de cada muestra, comenzando por la muestra de 100% de carne y 0% de grasa, concatenando seguidamente los espectros de la muestra de 95% de carne y 5% de grasa, y así sucesivamente con el resto de las muestras de diferentes porcentajes de carne hasta llegar a la muestra 0% carne y 100% tocino. Se utiliza también la referencia de grasa para entrenar mejor al sistema de clasificación con toda la información posible. El vector de coeficientes contiene el porcentaje de carne correspondiente a cada observación, por ejemplo, para todos los espectros de entrenamiento tomados de la muestra que contiene un 95% de carne, el coeficiente será 95.

Una vez creada la matriz de observaciones y el vector de coeficientes, el siguiente paso es determinar el número de componentes óptimo para que el algoritmo PLS pueda realizar correctamente las



estimaciones y clasificar las muestras medidas. El criterio usado para determinar el número de componentes es el criterio del 1%, esto significa que cuando el porcentaje de variación de un componente al siguiente es menor del 1% ya no es necesario su uso.

Conocido el número de componentes que se tiene que usar para clasificar correctamente esas observaciones, se vuelve a aplicar el algoritmo PLS con el número de componentes ajustado. Por último, se valida el método con la otra mitad de todas las muestras que se ha reservado para comprobar el grado de acierto de la clasificación del porcentaje de carne que realiza el método, a través de la obtención del coeficiente de correlación de los espectros reservados para la validación y los valores o información espectral que el método ha determinado que corresponden a cada porcentaje de carne tras el proceso de entrenamiento.

## PLS región VISNIR

### *Carne de novilla*

#### ○ Reflectancia $R(\lambda)$

Tomando los valores de reflectancia medidos y calibrados directamente por el sistema VIS-NIR, aplicando el ajuste por PLS se obtienen los siguientes resultados:

Número de componentes= 5

$$R^2 = 0.9290.$$

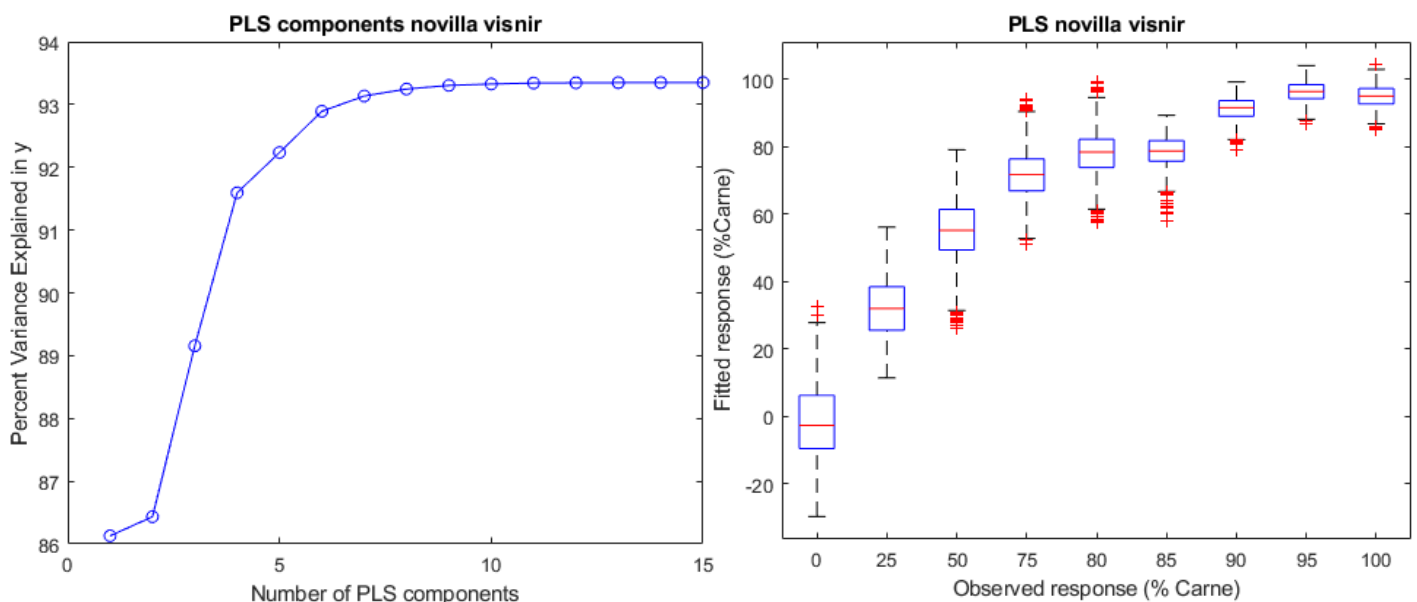


Figura 33: PLS en región VIS-NIR sobre carne de novilla

En la Figura 33 se observa el criterio del 1% a la hora de seleccionar el número de componentes mínimos necesarios para el correcto funcionamiento de estimación al

implementar PLS, en este caso el número de componentes es 5, ya que en el paso del componente 5 al 6 la pendiente varía menos de un 1%.

La figura se ha ayudado de la herramienta boxplot de la derecha para entender mejor la clasificación que realiza el método PLS. Cada una de las cajas de registro, o *boxplot*, permite determinar la mediana tanto en el caso de todas las estimaciones realizadas tanto para un porcentaje de carne concreto conocido (fase de entrenamiento), como para la clasificación que se realice para validar el ajuste. También muestra los valores comprendidos dentro de los percentiles 25 y 75, es decir la concentración de valores más típicos que toman las estimaciones realizadas. Como los coeficientes de referencia de una misma muestra usados en el proceso de entrenamiento. En este caso, el porcentaje de carne es el mismo para todas las observaciones de dicha muestra, sin usar boxplots sólo se puede distinguir el rango de valores entre los que varía la estimación realizada, pero no refleja en torno a que valores de todo ese rango existe una concentración mayor de las estimaciones realizadas.

La Figura 33 muestra que las muestras con mayor porcentaje de carne se ajustan mejor, los valores estimados tienen una menor varianza que las muestras con mayor porcentaje de grasa. Por ejemplo, una medida de un pixel de validación de la muestra que contiene un 25% de carne (conocido) mediante el ajuste PLS su estimación varía entre los valores 25% y 38% (percentil 25 y 75 respectivamente), siendo 31% de carne la mediana de las estimaciones, existiendo un total de 13 valores entre los que una muestra del 25% puede ser clasificada. Sin utilizar boxplot el valor mínimo es 11% y valor máximo 56%.

En cambio, si observamos que ocurre con la muestra de 95% de carne, el valor de la mediana se sitúa en 96% de carne y los percentiles 25 y 75 en los valores 94% y 98% de carne. Por tanto, para la muestra que contiene un 95% de carne, la variación de los valores más probables que pueden tomar las estimaciones es menor, 4 frente a los 13 valores entre los que varía la muestra de 25% de carne al contener más grasa. Esto se debe a que los espectros de las muestras que contienen más grasa pueden ser más cambiantes debido a efectos de *scattering* de la grasa que hacen que aumente su variabilidad.

Como se puede observar en las muestras con porcentajes de carne de 85% y 100% el percentil 75 de las estimaciones no llega a tomar esos valores, por lo que los valores más probables que pueden tomar las estimaciones no alcanzan el valor conocido de porcentaje de grasa que contiene la muestra, aunque se aproximan bastante como para realizar una clasificación lo suficientemente precisa.

#### ○ **Reflectancia $\ln(R(\lambda))$ y $\log(R(\lambda))$**

En este caso se aplica ajuste PLS a los espectros de las muestras con la información de la reflectancia expresada mediante su forma logarítmica. Se obtienen los siguientes resultados:  
Número de componentes= 4  $R^2 = 0.9281$ .

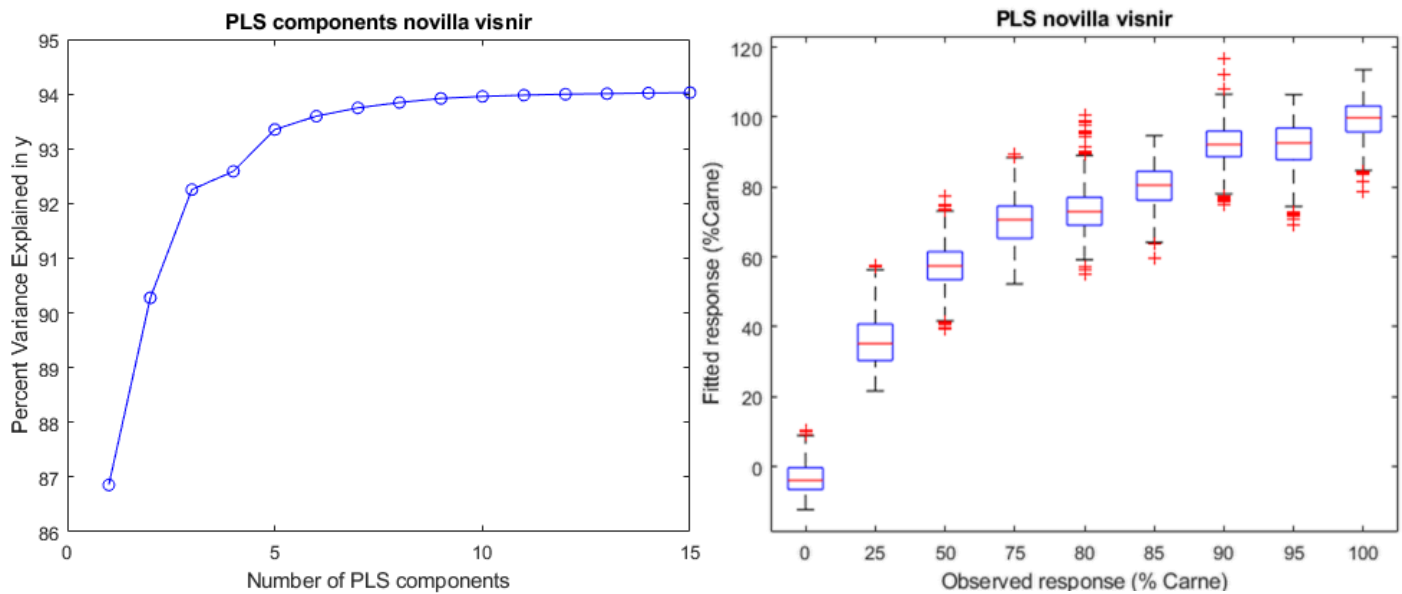


Figura 34: PLS en región VIS-NIR sobre carne de novilla (logaritmo)

En la Figura 34.a se observa que, dependiendo de cómo se representen los valores de reflectancia, aplicando o no el logaritmo de la reflectancia, el número de componentes PLS puede variar, necesitando 4 componentes para en el caso de  $\log(R(\lambda))$  y  $\ln(R(\lambda))$  y 5 cuando no se aplica logaritmo,  $R(\lambda)$ . No hay diferencias entre aplicar logaritmo natural o logaritmo decimal, se obtienen exactamente los mismos resultados.

En el caso de la muestra de 25% de novilla como se puede observar en la figura, el percentil 25 se sitúa por encima de ese valor en 31% por lo que la estimación da un porcentaje de carne mayor. En el caso de la muestra de 80% de carne ocurre lo contrario, los valores más probables que toma la estimación se quedan por debajo de ese valor, percentil 75 se sitúa en 77%.

Utilizando los valores de  $\log(R(\lambda))$  se observa un comportamiento similar al caso de la  $R(\lambda)$  en las muestras con menor porcentaje de carne, pero en cambio en las muestras con mayor porcentaje, véase una clara diferencia en la muestra de 95% de carne, aumenta la varianza de los valores más probables que pueden tomar la estimación. Por tanto, las muestras con mayor porcentaje de carne se ajustan mejor mediante PLS en el caso de  $R(\lambda)$ .

### Carne de Cerdo

Se sigue el mismo proceso que en la carne de novilla utilizando el sistema VIS-NIR y se examinan los resultados para los dos casos según como se representen los valores de la reflectancia.

- **Reflectancia  $R(\lambda)$**

Número de componentes= 5

$$R^2 = 0.9416$$

- **Reflectancia  $\ln(R(\lambda))$  y  $\log(R(\lambda))$**

Número de componentes= 4

$$R^2 = 0.9414$$

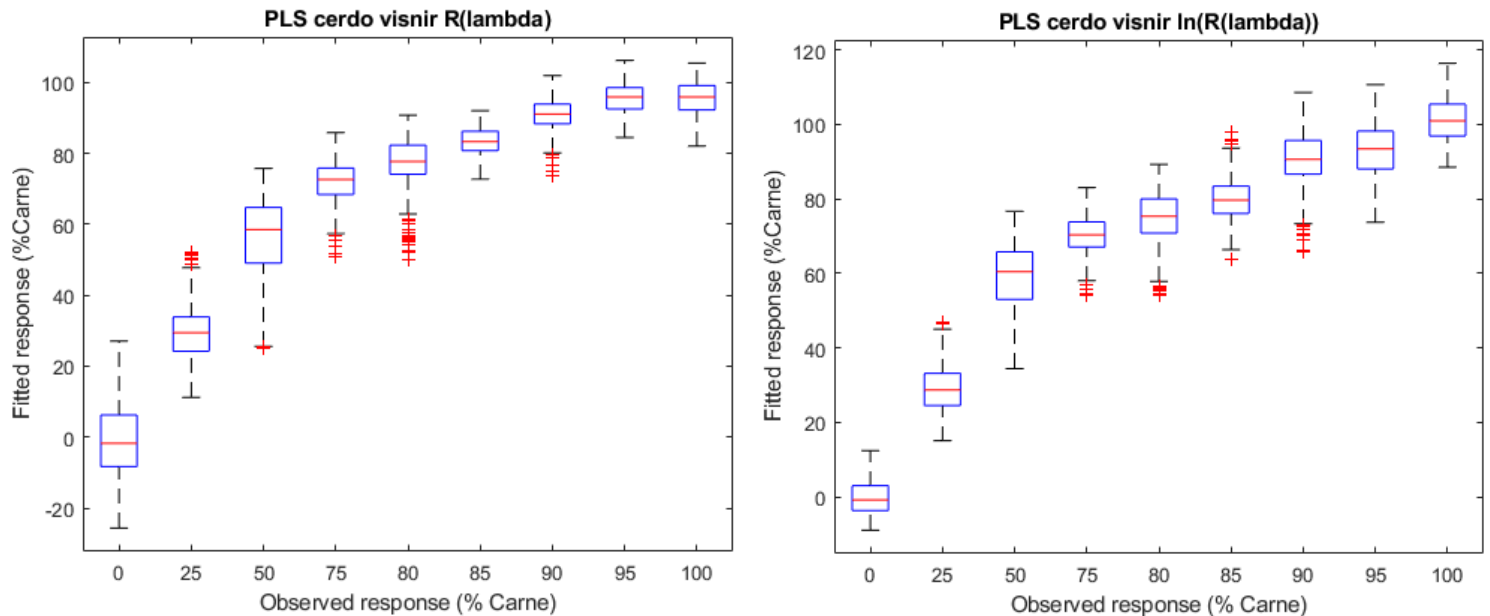


Figura 35: PLS en región VIS-NIR sobre carne de cerdo

### - Comparación PLS con $R(\lambda)$ de cerdo vs PLS con $R(\lambda)$ de novilla

Comparando con los resultados obtenidos para la carne de novilla en el supuesto reflectancia  $R(\lambda)$ , se observa que para las hamburguesas de cerdo en las muestras que presentan un mayor porcentaje de carne, la variabilidad de las estimaciones es mayor que en el caso de las hamburguesas de novilla. Esto se puede deber al hecho de que la carne de cerdo que se considera como pura (100% carne) contiene una mayor concentración de grasa que la referencia de 100% de carne de novilla.

Pero el coeficiente de correlación es más alto en el caso de la carne de cerdo. Aunque las muestras con menos porcentaje de grasa tengan menos variabilidad en el caso de la novilla, en las muestras con mayor porcentaje de grasa de cerdo esa variabilidad se compensa y no es tan alta como en las muestras con más grasa de novilla. Ese podría ser el motivo por el que se obtiene en conjunto un coeficiente de correlación mayor para las muestras de cerdo.

### - Comparación PLS con $R(\lambda)$ de cerdo vs PLS con $\ln(R(\lambda))$ de cerdo

El coeficiente de correlación obtenido en el caso de la reflectancia  $\ln(R(\lambda))$  es ligeramente menor que para  $R(\lambda)$ , como se puede apreciar en la Figura 35, las diferencias más significativas entre ambas gráficas se encuentran en las muestras con mayor porcentaje de carne donde los valores

estimados presentan una variación mayor en el caso de  $\ln(R(\lambda))$ , lo que podía provocar obtener menor coeficiente de correlación.

## PLS región SWIR

### *Carne de novilla*

Se sigue el mismo proceso que utilizando el sistema VIS-NIR y se examinan los resultados para los dos casos según como se representen los valores de la reflectancia.

- **Reflectancia  $R(\lambda)$**

Número de componentes= 5

$$R^2 = 0.9666$$

- **Reflectancia  $\ln(R(\lambda))$  y  $\log(R(\lambda))$**

Número de componentes= 4

$$R^2 = 0.9666$$

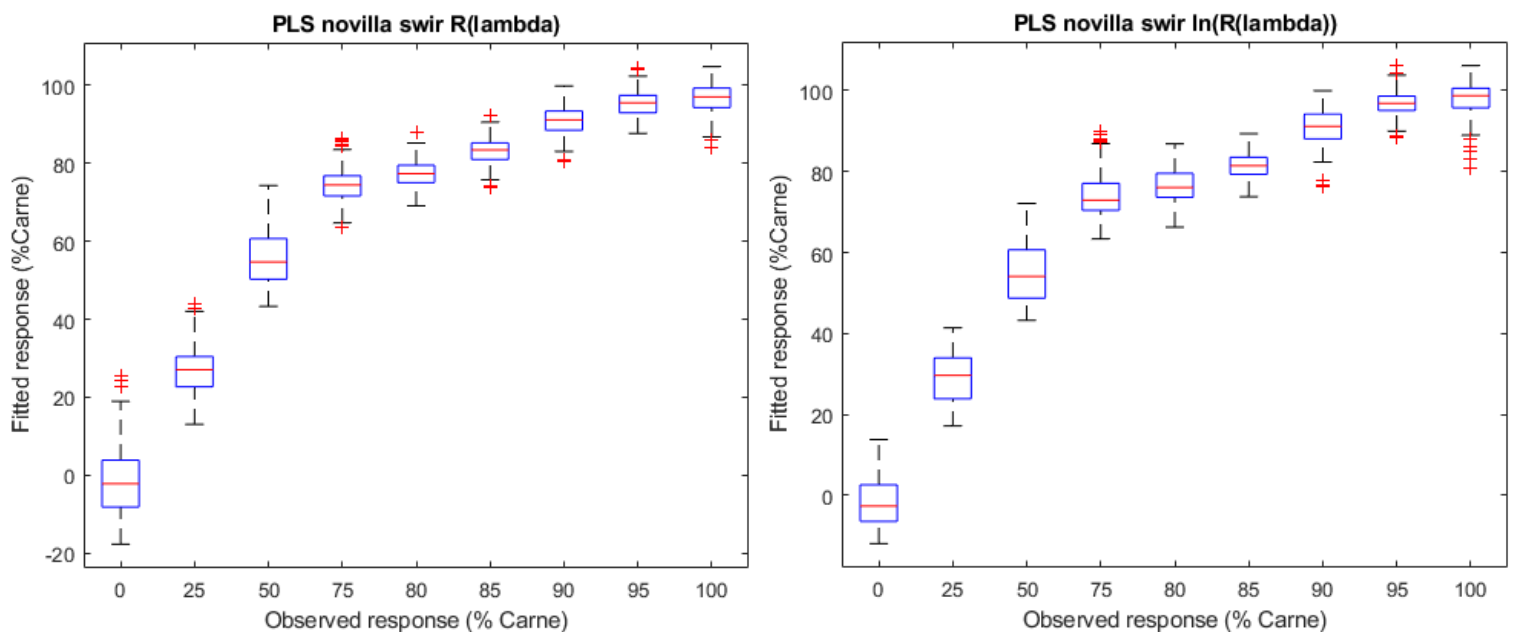


Figura 36: PLS en región SWIR sobre carne de novilla

### - Comparación PLS novilla SWIR vs novilla VIS-NIR

Independientemente de cómo esté expresada la reflectancia, midiendo con el sistema SWIR las muestras, se obtiene un mejor ajuste por PLS, con valores más altos que en el caso del sistema VIS-NIR, llegando a 0.966.

### - Comparación PLS con $R(\lambda)$ de novilla vs PLS con $\ln(R(\lambda))$ de novilla

Como se puede observar en la Figura 36, las gráficas con diferentes, pero se obtiene el mismo coeficiente de correlación para ambas expresiones de la reflectancia, tras la validación. Al usar los valores de reflectancia  $R(\lambda)$ , el ajuste PLS necesita 5 componentes para realizar las estimaciones. En cambio, en el caso de la reflectancia  $\ln(R(\lambda))$  se utilizan 4 componentes. Tras la fase de entrenamiento, los valores del coeficiente de correlación son diferentes en ambos casos, valiendo 0.9673 para reflectancia  $R(\lambda)$  y 0.9679 para  $\ln(R(\lambda))$ . Esa diferencia de número de componentes y los diferentes coeficientes de relación tras la fase de entrenamiento tienen como resultado que, después de validar el ajuste utilizando las muestras de validación, se obtenga el mismo coeficiente de correlación.

### Carne de cerdo

Se sigue el mismo proceso que en los casos anteriores y se examinan los resultados para los dos casos según como se representen los valores de la reflectancia.

- **Reflectancia  $R(\lambda)$**   
Número de componentes= 4  $R^2 = 0.9588$
- **Reflectancia  $\ln(R(\lambda))$  y Reflectancia  $\log(R(\lambda))$**   
Número de componentes= 4  $R^2 = 0.9657$

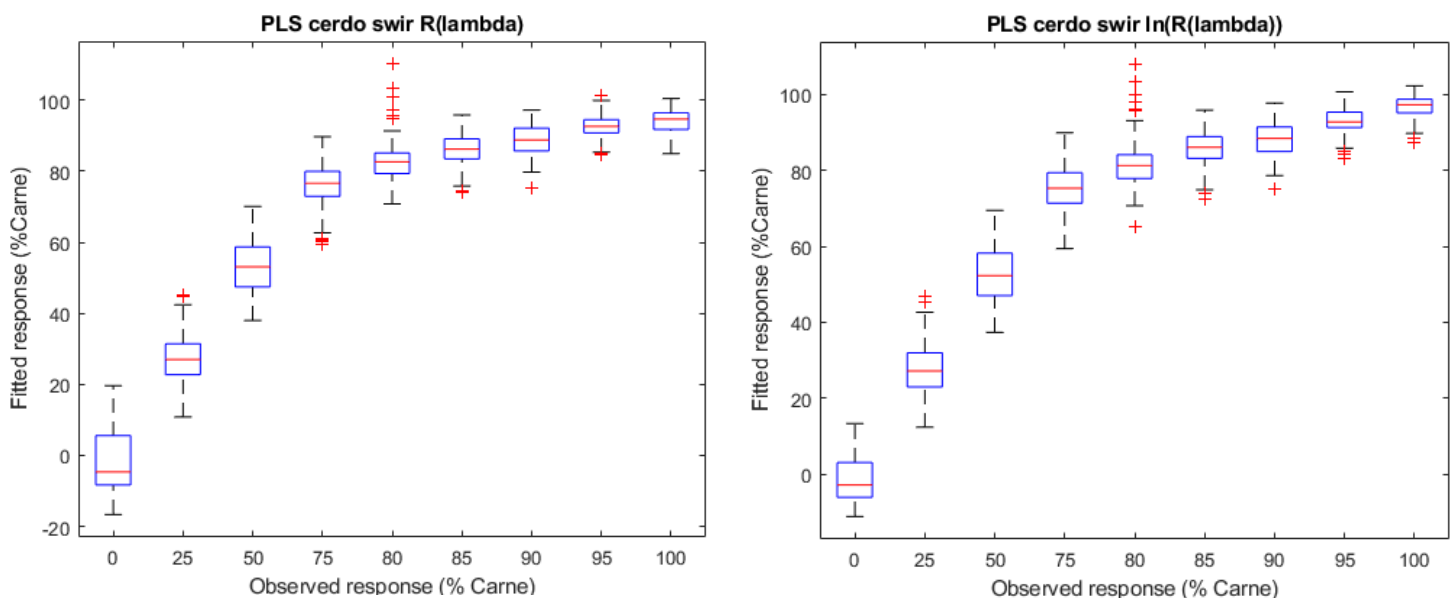


Figura 37: PLS en región SWIR sobre carne de cerdo

### - Comparación PLS cerdo SWIR vs cerdo VIS-NIR

Al igual que sucede con la novilla, la carne de cerdo medida con el sistema SWIR obtiene un coeficiente de correlación mayor que con el sistema VIS-NIR.

### - Comparación PLS con $R(\lambda)$ de cerdo vs PLS con $\ln(R(\lambda))$ de cerdo

En este caso, se obtiene un mejor coeficiente de correlación con la reflectancia del  $\ln(R(\lambda))$ . Puede ser debido a que, para la reflectancia de  $R(\lambda)$  se utilizan 4 componentes en lugar de 5 como en todos los casos anteriores.

Si se aumenta el número de componentes para  $R(\lambda)$  utilizando 5, el coeficiente de correlación mejora.  $R^2 = 0.9652$ , pero ya no se estaría cumpliendo el criterio del 1%, por lo que se perdería la referencia de comparación con el resto de los supuestos en los que se está aplicando el ajuste por PLS.

### PLS regiones VIS-NIR + SWIR

Después de haber analizado cada una de las regiones por separado, se quiere comprobar si el ajuste por PLS realiza una mejor estimación, es decir, obtiene un coeficiente de correlación más alto, si se tienen en cuenta todas las muestras de carne, o bien de cerdo o bien de novilla, pero en todo el rango del espectro, tanto con el sistema que realiza las medidas en la región VIS-NIR, como el que las realiza en SWIR.

#### *Carne de novilla*

Se sigue el mismo proceso que en los casos anteriores y se examinan los resultados para los dos casos según como se representen los valores de la reflectancia.

- **Reflectancia  $R(\lambda)$**   
Número de componentes= 7  $R^2 = 0.9185$
- **Reflectancia  $\ln(R(\lambda))$  y  $\log(R(\lambda))$**   
Número de componentes= 8  $R^2 = 0.9070$

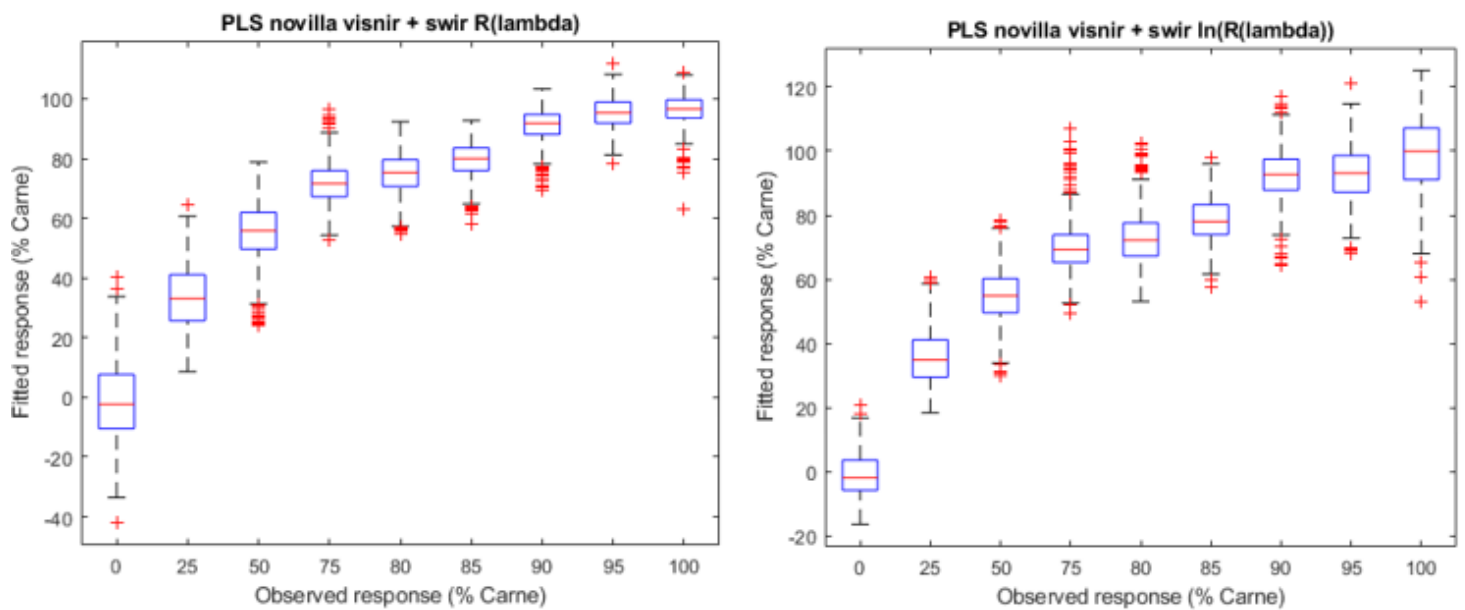


Figura 38: PLS en región VIS-NIR y SWIR sobre carne de novilla

Como la matriz de observaciones y el vector de coeficientes que contiene el porcentaje de carne para cada proporción es el doble en tamaño, al contener los datos de las dos regiones, se necesita un número de componentes más alto para realizar el ajuste. Como en la mayoría de los casos anteriores, la reflectancia  $R(\lambda)$  necesita un componente menor que para ajustar los valores de  $\ln(R(\lambda))$ . Vuelve a ser mayor el coeficiente de correlación para  $R(\lambda)$  que el caso de  $\ln(R(\lambda))$  como sucedía en el ajuste realizado para la novilla en VIS-NIR.

Al realizar el ajuste de las dos regiones, éste no mejora, sino que se obtiene un coeficiente de correlación más bajo. Esto puede ser porque los espectros de la zona VIS-NIR, comparándolos con los de la zona SWIR, son muy dispares en una región y en otra.

### Carne de cerdo

Se sigue el mismo proceso que en los casos anteriores y se examinan los resultados para los dos casos según como se representen los valores de la reflectancia.

- **Reflectancia  $R(\lambda)$**   
Número de componentes= 6  $R^2= 0.9324$
- **Reflectancia  $\ln(R(\lambda))$  y Reflectancia  $\log(R(\lambda))$**   
Número de componentes= 5  $R^2= 0.9237$



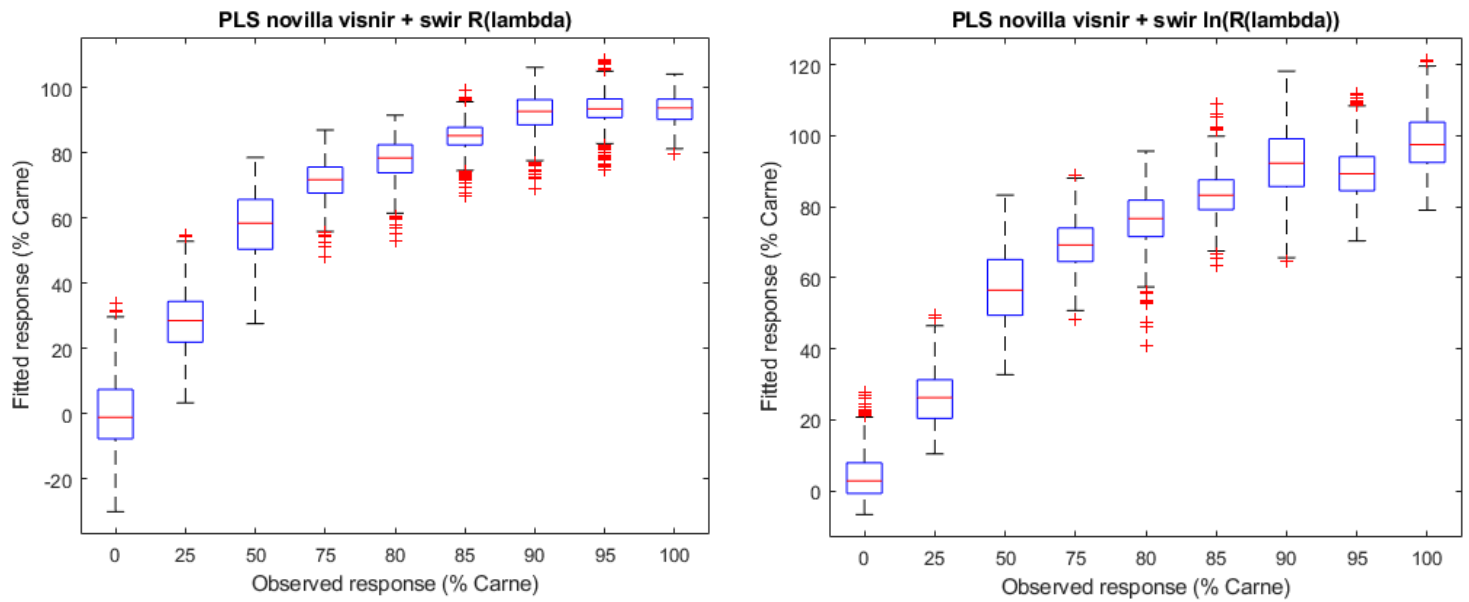


Figura 39: PLS en región VIS-NIR y SWIR sobre carne de cerdo

Juntando ambas regiones, para obtener el coeficiente de correlación de la carne de cerdo en su totalidad, no se obtienen mejores resultados que por separado, al igual que ocurre con la carne de novilla. La diferencia es que, en este caso, tanto para la reflectancia  $R(\lambda)$  como para  $\ln(R(\lambda))$  sólo que se añade 1 componente más a la hora de realizar el ajuste, juntando todos los datos. Comparándolo con la novilla, el coeficiente de correlación es más alto juntando ambas regiones del espectro, y los valores que toma para cada caso de reflectancia no son mucho menores respecto a los que se obtienen realizando el ajuste de PLS solamente en la región VIS-NIR.

Es mayor el coeficiente de correlación para  $R(\lambda)$  que el caso de  $\ln(R(\lambda))$  como sucedía en el ajuste realizado para la novilla en VIS-NIR.

### Visualización conjunta de resultados

Todas estas comparaciones que se han ido realizando anteriormente, quedan resumidas en la siguiente tabla, a través de los coeficientes de correlación para cada uno de los casos que se han contemplado.

PLS	$R_{test}^2$	VIS-NIR	SWIR	VIS-NIR + SWIR
Novilla	$R(\lambda)$	0.9290	0.9666	0.9185
	$\ln(R(\lambda))$	0.9281	0.9666	0.9070
Cerdo	$R(\lambda)$	0.9416	0.9588	0.9324
	$\ln(R(\lambda))$	0.9414	0.9657	0.9237

Tabla 11: Comparativa conjunta PLS en carne de novilla y cerdo

# Capítulo 7

## *Clasificación cualitativa*

Este capítulo tiene como objetivo analizar los resultados obtenidos utilizando las diferentes técnicas de procesado para la clasificación cualitativa explicadas en el capítulo 5, con las que se pretende estimar el porcentaje de grasa de cualquier hamburguesa a partir de las referencias de carne, tanto de cerdo como de novilla, que en este caso serán las muestras medidas con una concentración del 100% de carne y 0% de grasa añadida, y la referencia de grasa sería la muestra medida del 100% de tocino.

Se busca caracterizar el porcentaje de grasa y carne en una mezcla de la que a priori no se conoce qué porcentajes tiene, a través de las distintas técnicas de procesado como son la segmentación por distancia euclídea, SAM (*Spectral Angle Mapper*) y segmentación aplicando el algoritmo Otsu.

### 7.1 Clasificación por medio de OTSU

El primer paso para implementar la clasificación a través del algoritmo de Otsu es transformar la información tridimensional de la imagen hiperespectral con la información espacial y espectral de la muestra a una matriz bidimensional. Esta matriz contiene, para cada pixel de la hamburguesa un único dato representativo del espectro. Sobre esta matriz será sobre la que se aplique Otsu.

Se van a utilizar 3 formas distintas de obtener ese dato representativo del espectro. La primera es calculando el promedio de toda la información espectral para cada punto de la hamburguesa. La segunda es seleccionando el valor máximo del espectro (método MIP, *Maximum Intensity Projection*), es decir, la mayor reflectancia; y la tercera el valor mínimo, que es el método invMIP.

Se calcula el promedio, el MIP y el invMIP para cada tipo de carne.

Sobre la matriz resultante de cada método, se aplica el algoritmo Otsu y se visualiza la máscara para tener una información espacial de lo que el algoritmo clasifica como grasa y lo que clasifica como carne, además de obtener el histograma para ver la distribución de los valores de reflectancia en cada caso y ver si el nivel que establece el algoritmo se corresponde con esa distribución.

### 7.1.1 Carne de novilla

#### Región VIS-NIR

Una vez aplicado el algoritmo de Otsu sobre el promedio, MIP e invMIP, se obtienen los diferentes umbrales en los que todos los valores que estén por encima de ese nivel se clasifican como grasa, y por debajo, como carne. Dependiendo del método que se utilice, ese umbral tomará un valor u otro. Los valores por encima de ese umbral se clasifican como grasa porque se corresponden con una mayor reflectancia, y la grasa refleja más que la carne.

Umbrales de Otsu que se obtienen para la carne de novilla en la región VIS-NIR:

- Promedio: 0.1922
- MIP: 0.2863
- invMIP: 0.0863

Observando la distribución de los histogramas de la Figura 40 y los umbrales para cada método, se evidencia una correlación, comprobando que el algoritmo Otsu establece bien el límite entre lo que clasifica como grasa o carne.

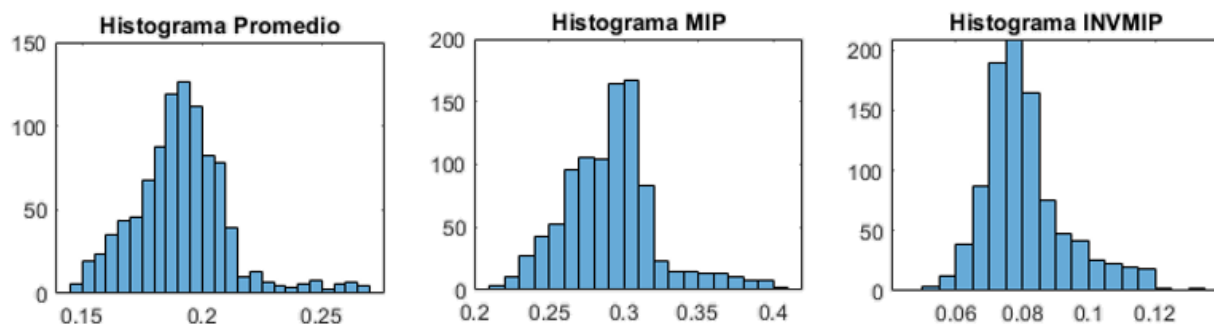


Figura 40: Histogramas del algoritmo Otsu sobre carne de novilla en VIS-NIR

En la Figura 41 se observan las máscaras o zonas de la muestra que cada método identifica (representadas en blanco) como grasa o como carne. El Otsu del método MIP se aplica sobre los valores máximos de reflectancia del espectro en cada punto de la muestra, es lógico que el umbral sea alto y como se puede comprobar en la imagen, en consecuencia, este método potencia la grasa.

Al contrario que ocurre con el Otsu del método invMIP que está formado por los valores mínimos espectrales, y por tanto el umbral es más bajo, potencia la clasificación de más región de la muestra como carne. Y el promedio, establece un umbral que toma un valor medio entre los otros dos casos, hecho que también aprecia en la representación de las zonas que distingue como grasa y carne.

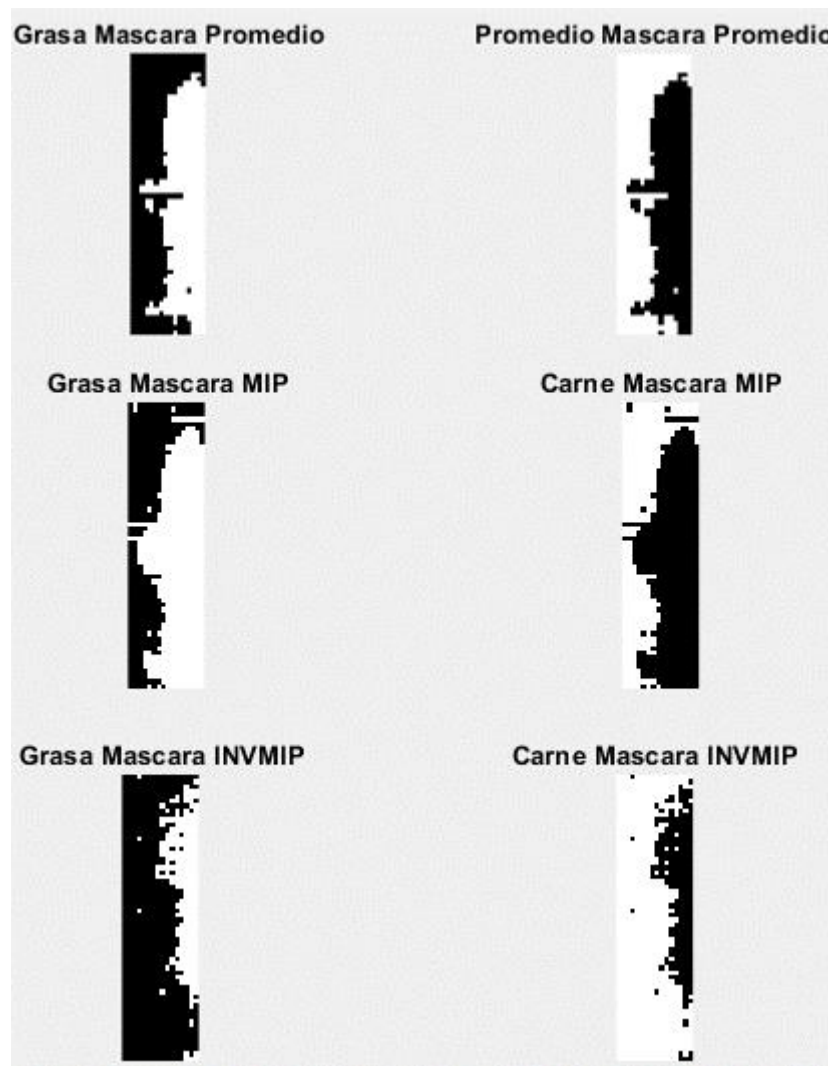
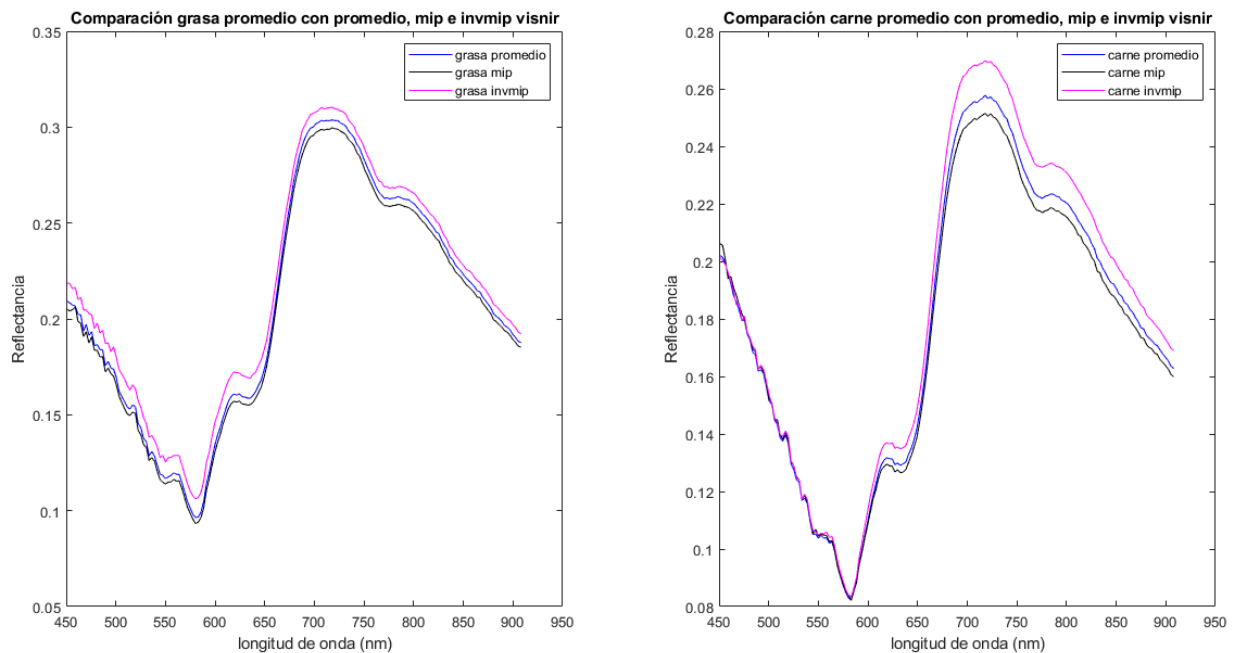


Figura 41: Máscaras en carne y grasa de otsu en novilla VIS-NIR

Para poder comprobar la eficacia de la clasificación a través del algoritmo Otsu aplicado bien al promedio, al MIP o al invMIP, es necesario correlar la grasa de referencia (100% tocino) con lo que Otsu clasifica como grasa en cada uno de los tres casos. Utilizando la máscara obtenida mediante Otsu correspondiente a cada uno de los tres métodos (promedio, MIP, invMIP), se reproduce esa matriz con la máscara a un cubo con las mismas dimensiones que la imagen hiperespectral, de forma que la máscara se pueda aplicar a cada longitud de onda al comparar el cubo de la máscara con el cubo de espectros y obtener así toda la información espectral de los píxeles de la muestra que se habían clasificado anteriormente como carne o como grasa por medio de uno de los métodos utilizados (promedio, MIP, invMIP). Esto sirve para obtener por un lado los espectros de todos los puntos clasificados como grasa y calcular seguidamente el espectro promedio de la grasa. Y por otro lado, obtener el espectro promedio de la carne.

Este proceso se lleva a cabo para cada método en el que se ha basado Otsu para clasificar. Por tanto, obtendremos por un lado los espectros promedio de la grasa y de la carne del Otsu del promedio, los espectros promedio de la grasa y de la carne del Otsu del MIP y los espectros promedio de la grasa y de la carne del Otsu del invMIP. Todos estos espectros promedios se encuentran representados en la Figura 42.



**Figura 42:** Comparación grasa y carne novilla en VIS-NIR usando promedio, MIP e invMIP

Al observar los resultados de cualquiera de los tres métodos se cumple que la reflectancia de la grasa es mayor que la de la carne. Por lo tanto, se comprueba visualmente el funcionamiento en el que basa el algoritmo Otsu. No ocurre en caso de la grasa debido a que los espectros de los tres métodos toman valores muy similares, pero si en el de la carne, donde es más evidente y se cumple que el método de clasificación Otsu basado en invMIP, es decir con un umbral más bajo porque contenía los valores mínimos, de forma que potenciaba la clasificación de más región de la muestra como carne, tenga valores más altos de reflectancia para la carne ya que más muestras han sido clasificadas como carne según un único valor bajo de su espectro pero pudiendo tener en el resto de longitudes de onda valores de reflectancia altos y eso se refleja en el espectro.

Una vez obtenidos los tres espectros promedio de grasa, se comparan con la referencia conocida de grasa, la muestra de tocino 100%, para conocer así el grado de correlación de la grasa estimada en cada caso con la grasa pura conocida. Obteniendo así la certificación de la clasificación cualitativa.

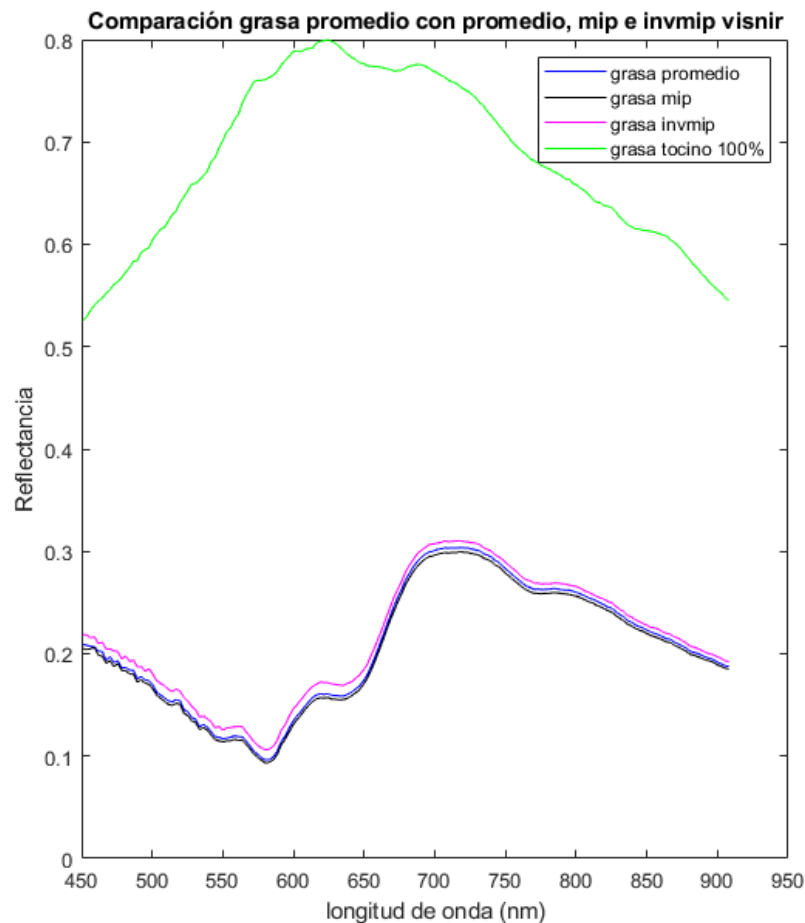


Figura 43: Comparación grasa novilla en VIS-NIR frente al valor de referencia

En la Figura 43 se aprecia que la forma del espectro de grasa de referencia además de tener valores de reflectancia mucho más elevados como cabía esperar, tiene una forma muy dispar a los espectros promedio de la muestra que mediante el Otsu de los diferentes métodos se han clasificado como grasa, debido a que los espectros de grasa estimados se ven muy influenciados por la carne, puesto que en la muestra de referencia de carne que se está analizando es el componente principal. Esta diferencia significativa en la forma de los espectros apreciable de forma visual también queda evidenciada a través de la correlación con la grasa de referencia.

$r(x,y)$	Promedio	MIP	INMIP
Referencia (T 100%)	-0.0027	-0.0075	0.0112
Promedio	1	1.0000	0.9996
MIP	1.0000	1	0.9994
INVMIP	0.9996	0.9994	1

Tabla 12: Comparativa novilla VIS-NIR promedio, MIP e INVMIP

En la tabla también se representa la correlación entre los distintos valores del espectro de grasa estimados con Otsu del promedio, MIP e invMIP. Toman valores muy similares por lo que la correlación entre ellos es muy alta, llegando incluso al valor  $r=1$  cuando se comparan el promedio y el MIP.

### Región SWIR

En el caso de las medidas de carne de novilla realizadas con el sistema SWIR, el proceso para clasificar cualitativamente mediante Otsu el porcentaje de grasa es el mismo que el explicado previamente para el sistema VIS-NIR, por lo que se expondrán los resultados obtenidos para poder realizar comparaciones.

Umrales del algoritmo Otsu que se obtienen para la carne de novilla en la región SWIR:

- Promedio: 0.2706
- MIP: 0.3490
- INVMIP: 0.2000

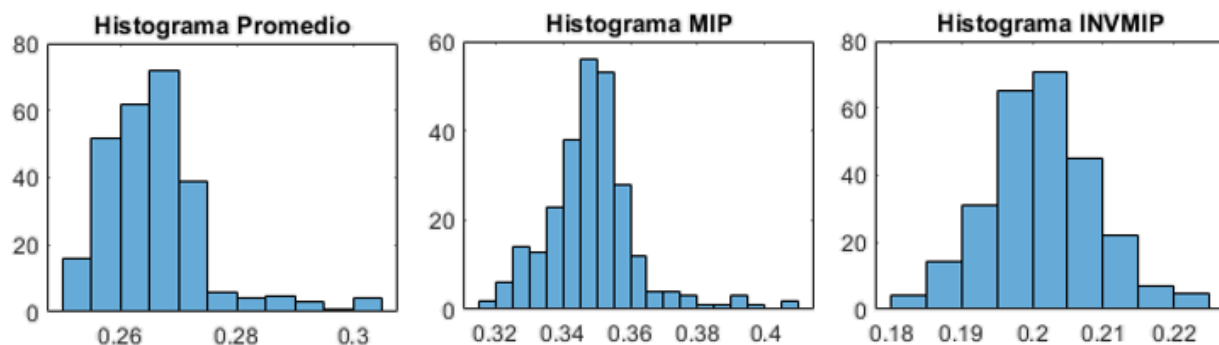


Figura 44: Histogramas del algoritmo Otsu sobre carne de novilla en SWIR.

Observando los umbrales y comparándolos con los histogramas, se comprueba que los umbrales de Otsu están correctamente escogidos por el algoritmo. También, observando los histogramas se puede entender las representaciones de las máscaras, y cómo en el caso del invMIP el umbral se sitúa en el medio, asemejándose a una distribución gaussiana, por lo que existe el mismo nivel de valores considerados como carne que como grasa. En el caso del promedio, a la derecha de umbral fijado, que sería la grasa, están contenidos muchos menos píxeles que a la izquierda, por la distribución que adopta, clasificándose muchos más píxeles con los valores promedio como carne. En la imagen se aprecia dentro de la región de la ROI (*Region of Interest*) que se ha seleccionado de cada muestra medida por el sistema SWIR, lo que el algoritmo Otsu clasifica como carne y grasa.

Al igual que ocurría en la región VIS-NIR, cuando se utiliza el MIP se potencia la grasa, pero en cambio en este caso la máscara de promedio clasifica más región de la muestra como carne que con la máscara invMIP por el motivo explicado previamente.

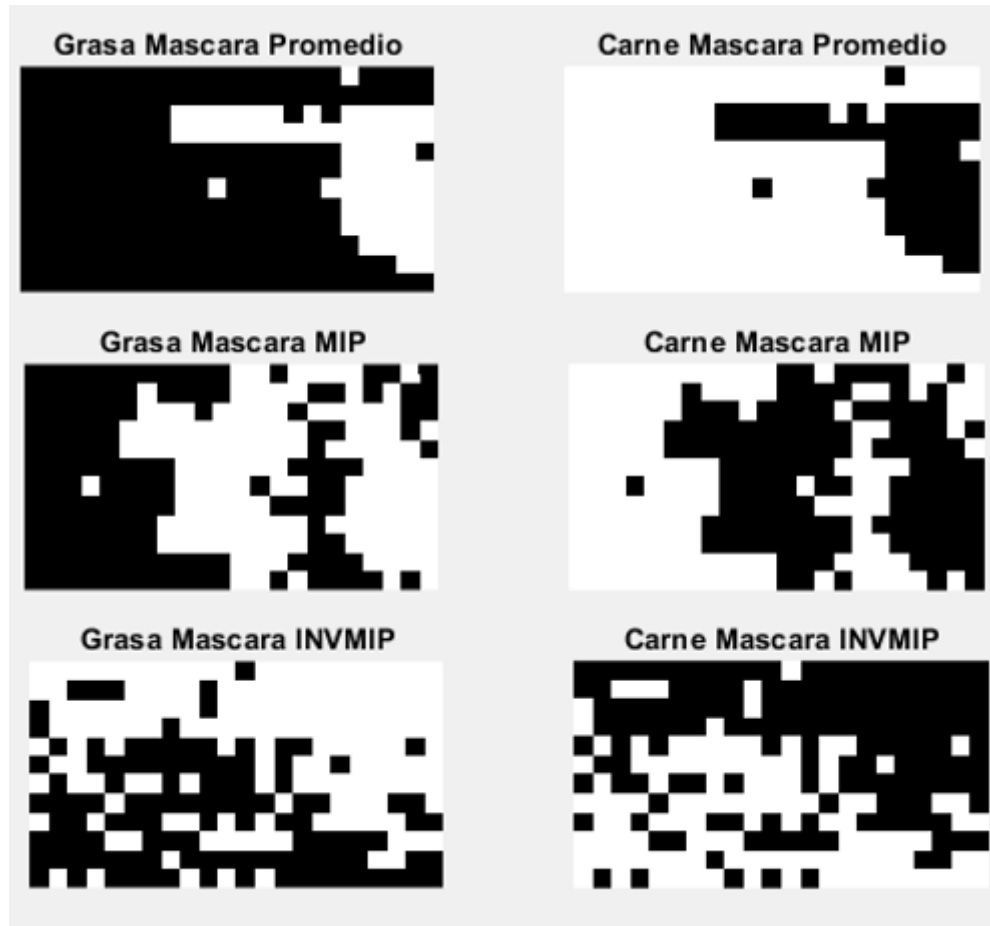


Figura 45: Máscaras en carne y grasa de otsu en novilla SWIR

Como se ha explicado en el caso del sistema VIS-NIR, y por cómo funciona el algoritmo Otsu, ya se ha comprobado que los espectros promedio de los tres métodos siempre van a tener valores más altos de reflectancia que la carne, por lo que se va a analizar directamente la correlación entre las grasas estimadas y la de referencia. En la representación de estos espectros promedio de grasa, en la Figura 46, se aprecia que ahora en la región SWIR la forma del espectro de grasa de referencia tocino 100% correla mejor con la forma de los espectros de los puntos de la muestra clasificados como grasa, con el MIP, el invMIP y el promedio. Los valores de correlación se pueden apreciar en la Tabla 14. Se puede observar que, en el caso de la comparación del espectro de grasa promedio con el de referencia, se obtiene un coeficiente de correlación del 0.9, y en torno a 0.89 para los otros dos casos. Entre los distintos métodos, la correlación toma valores próximos a 1.



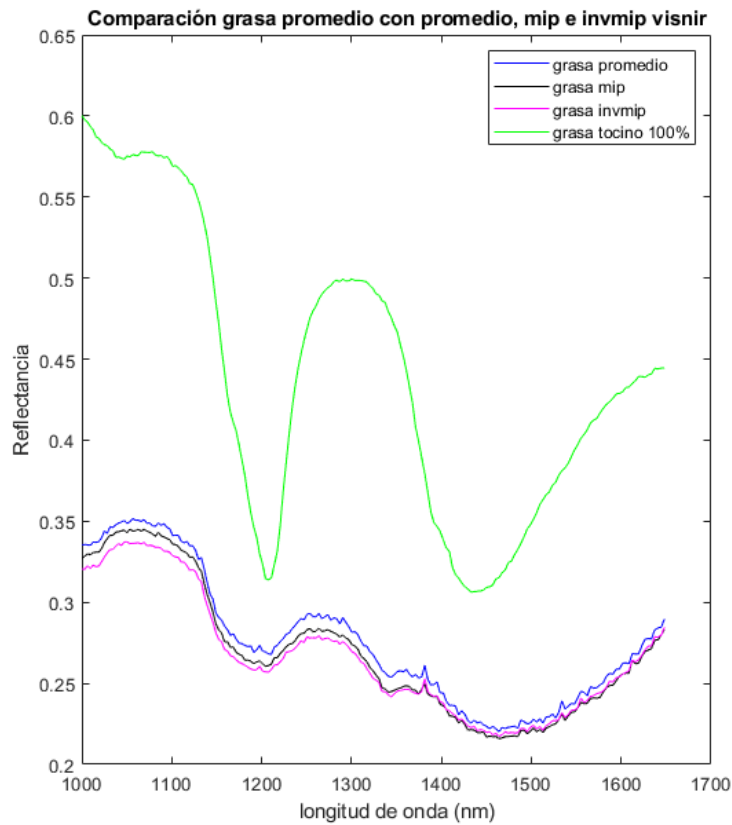


Figura 46: Comparación grasa novilla en SWIR frente al valor de referencia

$r(x,y)$	Promedio	MIP	INMIP
Referencia (T 100%)	0.9030	0.8992	0.8993
Promedio	1	0.9992	0.9974
MIP	0.9992	1	0.9989
INVMIP	0.9974	0.9989	1

Tabla 13: Comparativa novilla SWIR promedio, MIP e INVMIP

### 7.1.2 Carne de cerdo

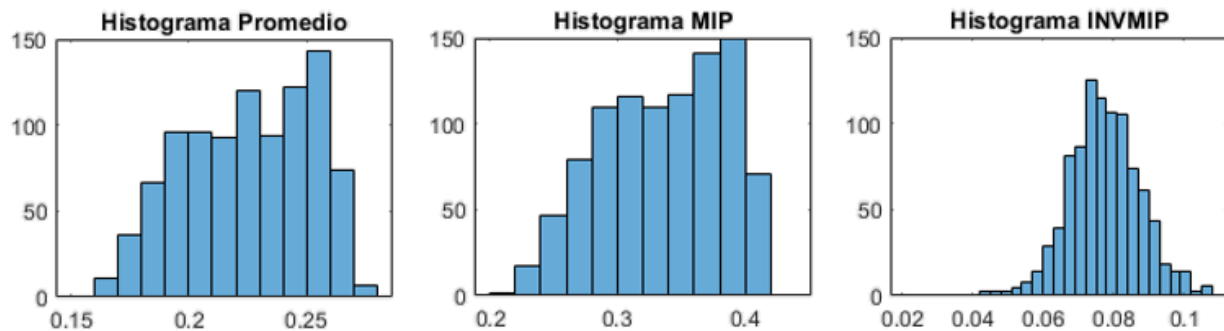
Se repite el mismo proceso seguido para la carne de novilla, por lo que se obtendrán directamente los resultados, pudiendo así establecer distintas comparaciones.

#### *Región VIS-NIR*

Umbral del algoritmo Otsu que se obtiene para la carne de cerdo en la región VIS-NIR:

- Promedio: 0.2196
- MIP: 0.3294
- invMIP: 0.0745

La forma de los histogramas para la carne de cerdo en VIS-NIR son muy parecidos para el promedio y el invMIP, aunque el rango de valores cambia.



**Figura 47: Histogramas del algoritmo otsu sobre carne de cerdo en VIS-NIR.**

Comparando los umbrales obtenidos por el sistema VIS-NIR para la carne de novilla y la carne de cerdo, se puede concluir que toman valores parecidos. Observando las máscaras, la distribución de los histogramas y dónde se sitúa el umbral dentro de esa distribución, se ve que en el caso del promedio y el MIP las distribuciones son parecidas con el umbral en el medio, por lo que en torno al 50% de la muestra se considera grasa, y el otro 50% carne. Y en el caso del invMIP, aunque se tomen los valores mínimos y se debería potenciar la carne, por la forma de la distribución, hay muchos más píxeles por encima del umbral, por lo que la máscara se refleja como grasa.

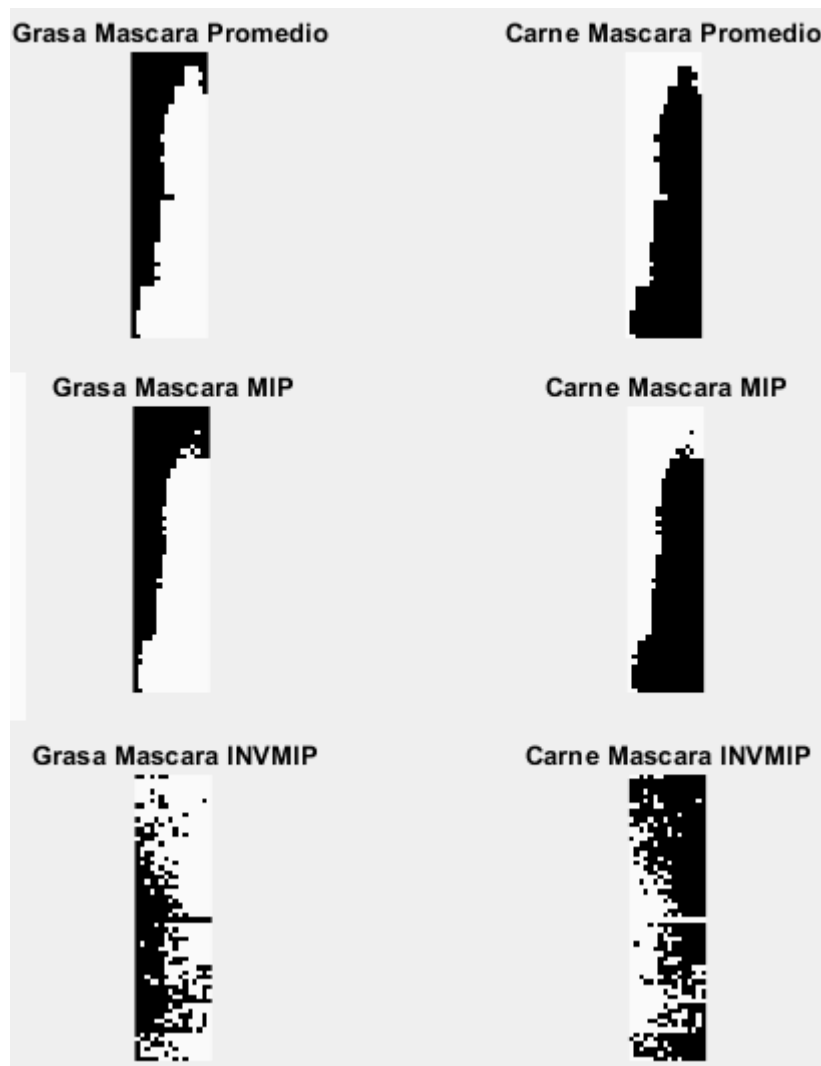


Figura 48: Máscaras en carne y grasa de otsu en cerdo VIS-NIR

En la gráfica donde se encuentran los espectros de las grasas de referencia y las grasas estimadas representados, para la carne de cerdo, se observa que los espectros clasificados como grasa tienen valores de reflectancia más altos que para los espectros clasificados como grasa para la carne de novilla (Figura 49). Esto coincide con discusiones realizadas anteriormente relacionadas con que la carne de cerdo contiene un porcentaje de grasa intrínseco mayor.

Aunque la forma del espectro de la grasa de referencia en comparación con los otros tres sigue siendo muy dispar, existe una correlación un poco más alta que en el caso de la carne de novilla, porque la forma de estos tres espectros no está tan influenciada por la carne al tener de por sí el cerdo una concentración de grasa mayor.

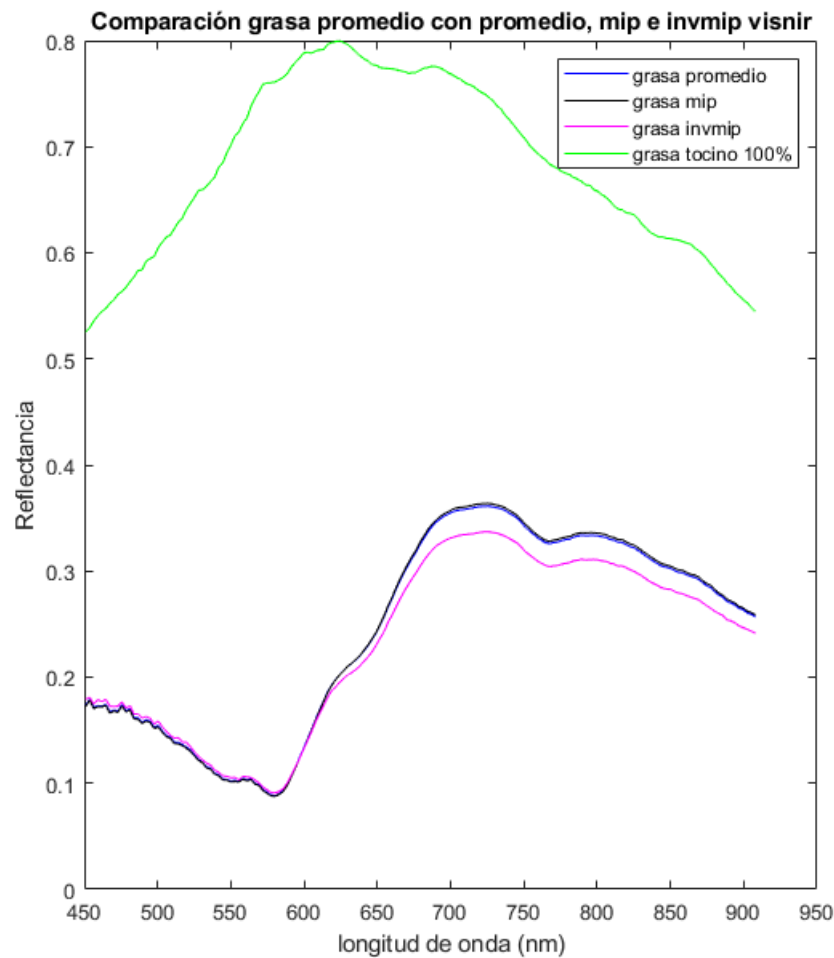


Figura 49: Comparación grasa cerdo en VIS-NIR frente al valor de referencia

Aunque la forma del espectro de la grasa de referencia en comparación con los otros tres sigue siendo muy dispar, existirá una correlación un poco más alta que en el caso de la carne de novilla, porque la forma de estos tres espectros muy similares entre ellos no está tan influenciada por la carne al tener de por sí el cerdo una concentración de grasa mayor. Los valores de correlación se pueden observar en la tabla. Como los espectros de grasa del MIP y el promedio son tan parecidos, tienen la misma correlación al compararlos con la grasa de referencia.

$r(x,y)$	Promedio	MIP	INMIP
Referencia (T 100%)	0.0409	0.0409	0.0282
Promedio	1	1.0000	0.9994
MIP	1.0000	1	0.9993
INVMIP	0.9994	0.9993	1

Tabla 14: Comparativa cerdo VISNIR promedio, MIP e INVMIP

### *Región SWIR*

Umbrales del algoritmo Otsu que se obtiene para la carne de cerdo en la región SWIR:

- Promedio: 0.2667
- MIP: 0.3961
- INVMIP: 0.1765

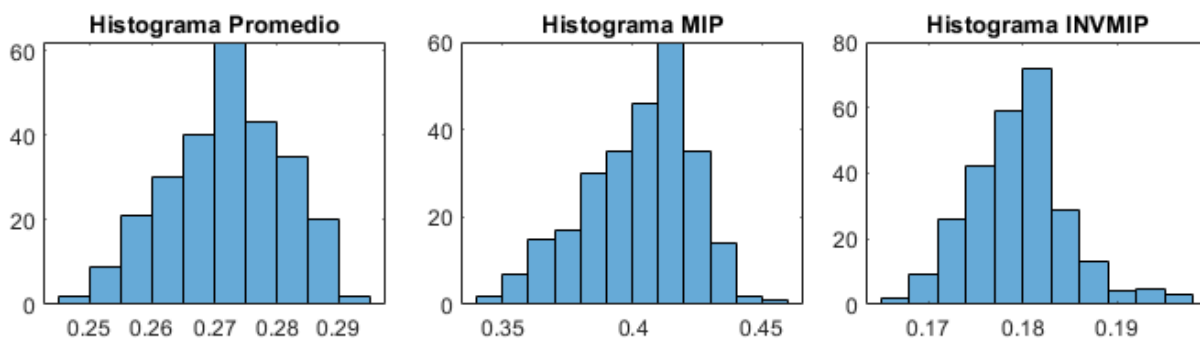


Figura 50: Histogramas del algoritmo otsu sobre carne de cerdo en SWIR.

Tras comparar como en los casos anteriores las máscaras con la distribución de los histogramas y el umbral definido por el algoritmo Otsu, se puede concluir que, en este caso, la máscara del MIP

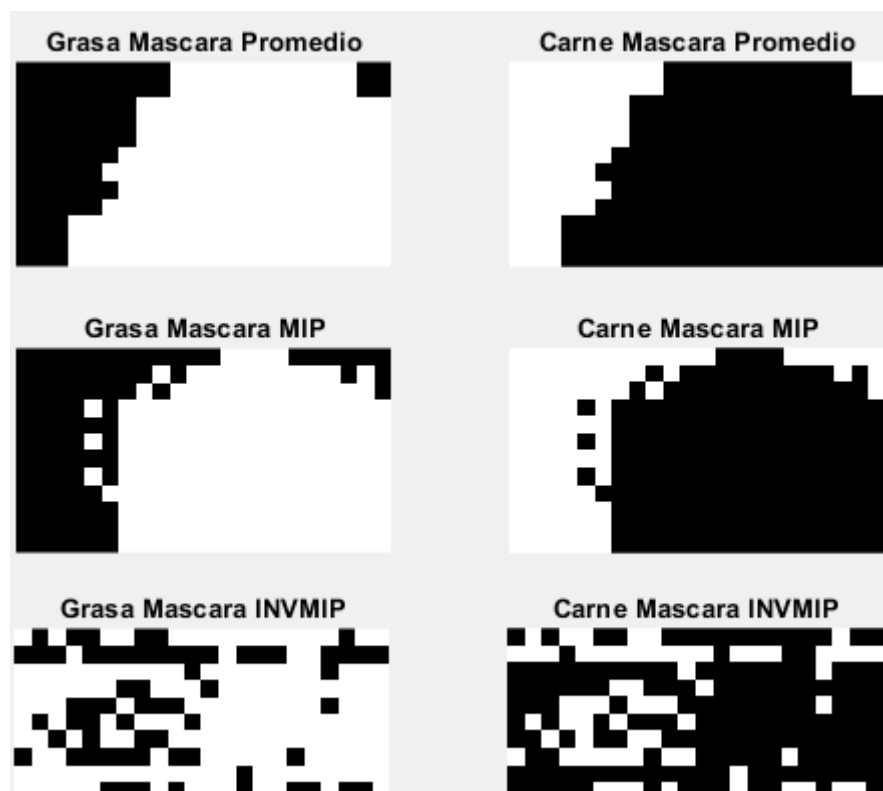


Figura 51: Máscaras en carne y grasa de otsu en cerdo SWIR

y el promedio vuelven a ser muy parecidas, y en este caso el invMIP clasifica más región de la hamburguesa como grasa.

Obteniendo los espectros de las grasas, visualmente se puede observar en la Figura 52 que, como en el caso de la novilla, en la región SWIR hay una correlación más alta con la grasa de referencia, cuyos valores se representan en la Tabla 16.

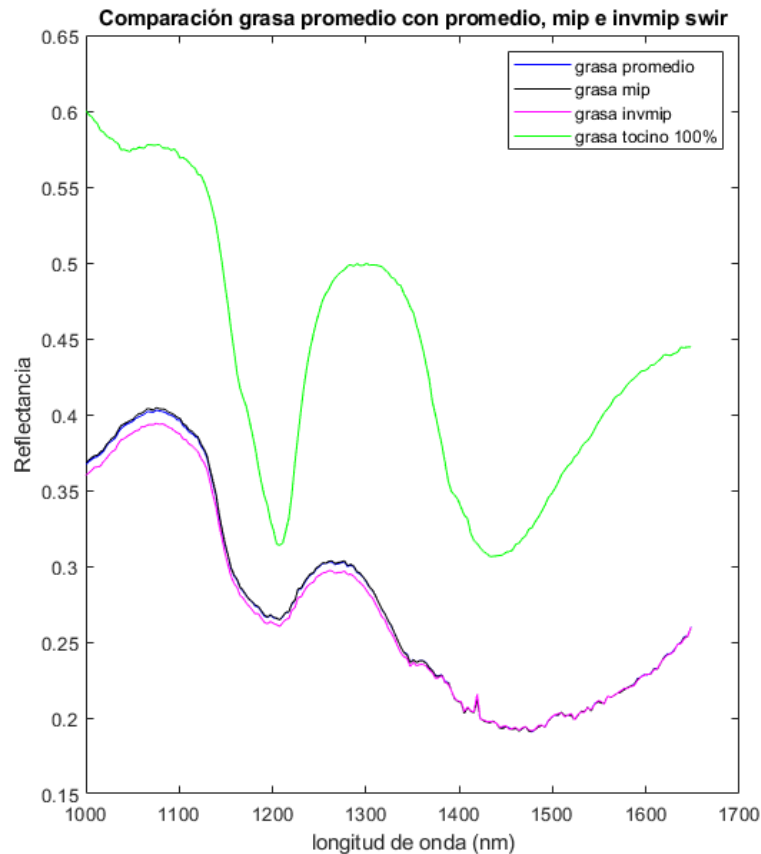


Figura 52: Comparación grasa cerdo en SWIR frente al valor de referencia

$r(x,y)$	Promedio	MIP	INMIP
Referencia (T 100%)	0.8904	0.8899	0.8914
Promedio	1	1.0000	0.9999
MIP	1.0000	1	0.9999
INVMIP	0.9999	0.9999	1

Tabla 15: Comparativa cerdo SWIR promedio, MIP e INVMIP



Al comparar la correlación de las grasas estimadas con la de referencia, se obtienen valores un poco inferiores a los obtenidos para la carne de novilla. Las correlaciones entre los tres métodos son prácticamente todas máximas, iguales a 1.

## 7.2 Clasificación mediante distancia euclídea

Este método, para realizar una clasificación, necesita tener un prototipo de grasa como referencia a partir del cual se calcula la distancia de separación espectral. En este caso se utilizará el espectro promedio de la muestra de 100% de tocino, y el prototipo de carne de referencia serán las muestras medidas de 100% de novilla y 100% de cerdo según el caso. Esta clasificación consiste en recorrer pixel por pixel la muestra que se quiere clasificar, obteniendo para cada punto toda la información espectral. El siguiente paso es comparar el espectro de ese punto por un lado con el espectro de referencia de grasa a través de la distancia euclídea de ambos vectores. Si la distancia es pequeña quiere decir que se asemeja bastante a la grasa y se clasifica como tal, mientras que si la distancia es grande estaría más próximo a la carne. Cuando se compara ese espectro en ese punto con la referencia de carne, es el mismo razonamiento, cuando más pequeña la distancia más certeza habrá de que el material a evaluar es carne, y si la distancia es grande, es grasa.

Lo que se pretende es obtener el valor de la distancia euclídea para cada punto de la muestra al compararlo con la referencia de grasa y, con el valor máximo y mínimo obtenidos, representar la distancia a través de una escala de colores degradada en función de esos valores localizando visualmente de esta manera las regiones donde se concentra más grasa o más carne. Se realiza otra representación del mismo modo en este caso de la distancia euclídea de cada punto de la muestra con la referencia de carne también a través de una escala de colores. El proceso que se repite para cada una de las muestras medidas de ese tipo de carne, es decir, donde ya se conocen los porcentajes de grasa que contienen dichas las mezclas, desde la mezcla 100% de carne y 0% de grasa, hasta la muestra 25% de carne y 75% de grasa.

Juntando todos los resultados en una única representación gráfica se espera observar un degradado que representa la variación de los porcentajes de grasa y de carne de las composiciones. Además de observar la distribución de ambos componentes en cada muestra, es decir el grado de homogeneidad que presenten.

### 7.2.1 Carne de novilla

Todo lo explicado anteriormente se va a aplicar sobre la carne de novilla, por un lado, utilizando el sistema VIS-NIR y por otro el SWIR.

#### *Región VIS-NIR*

Al comparar los espectros promedio de cada muestra con diferentes porcentajes de carne de novilla con los prototipos de grasa y carne de referencia, se veía que, en algunas zonas del espectro, la

diferencia euclídea era demasiado grande o demasiado pequeña en algunos puntos, como en los picos de la hemoglobina que se juntaban todos los espectros con el de referencia de la carne, existiendo una diferencia demasiado grande con la grasa como para poder comparar ambas representaciones con la misma escala de valores.

A la hora de representarlo gráficamente no se apreciaban las variaciones, no pudiendo así obtener resultados concluyentes. Por lo que se recorta el espectro (650 a 950nm) seleccionando las longitudes de onda en las que la forma de los espectros promedio de todas las muestras, dependiendo del porcentaje de novilla, presentan los cambios más significativos y de una manera lineal, como se puede observar en la Figura 53.

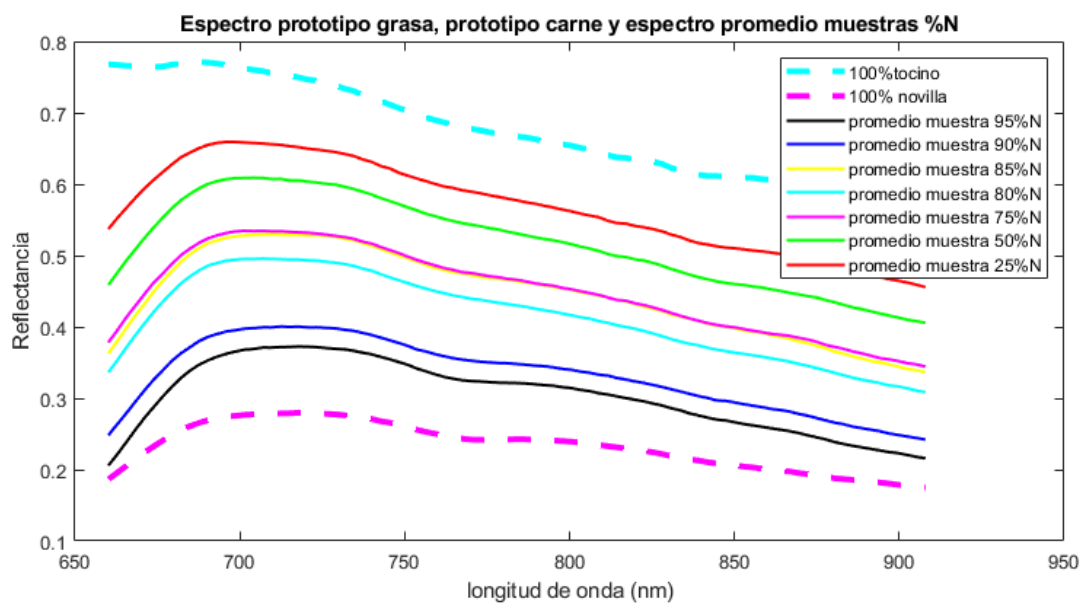


Figura 53: Espectros recortados en novilla VIS-NIR para distancia euclídea

Para poder representar bajo la misma escala de valores, tanto la distancia respecto a la grasa de referencia como a la carne, se determina el valor máximo, en este caso superior a 11 que se representa con el color rojo más oscuro, y el mínimo en azul oscuro. Las muestras de la primera fila de la figura, con el porcentaje de carne que se indica, se comparan con la referencia de carne de novilla, viendo que a medida que aumenta el porcentaje de grasa de las mezclas, también aumenta la distancia. En el caso de la segunda fila, se compara el porcentaje de grasa de las muestras con la grasa de referencia, y a medida que aumenta el porcentaje de grasa en las mezclas, la distancia al promedio del 100% de tocino es menor, reflejando ese aumento de la cantidad de grasa. Comparando las dos figuras, aunque en la de los espectros sólo se ha representado el promedio de cada muestra en lugar de representar todos los espectros de todos los puntos, sirve para hacerse una idea de lo que varían las distancias y lo que se espera obtener en la representación espacial.



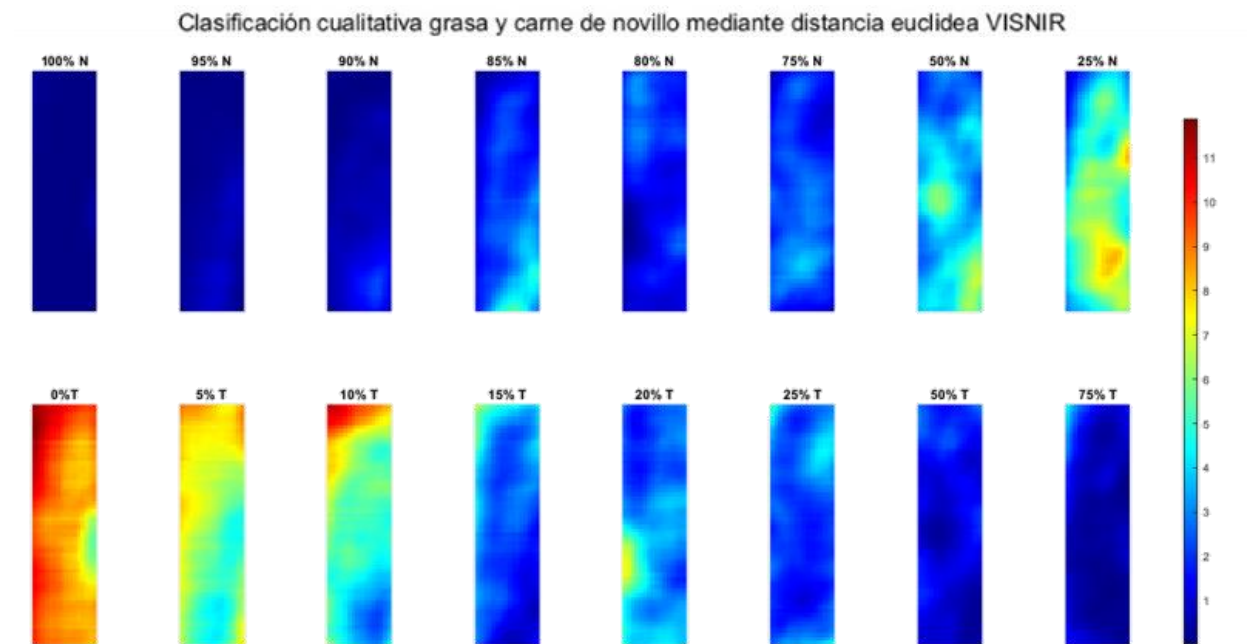


Figura 54: Clasificación cualitativa en novilla VIS-NIR mediante distancia euclídea

### Región SWIR

Se muestran ahora los resultados para el rango espectral SWIR.

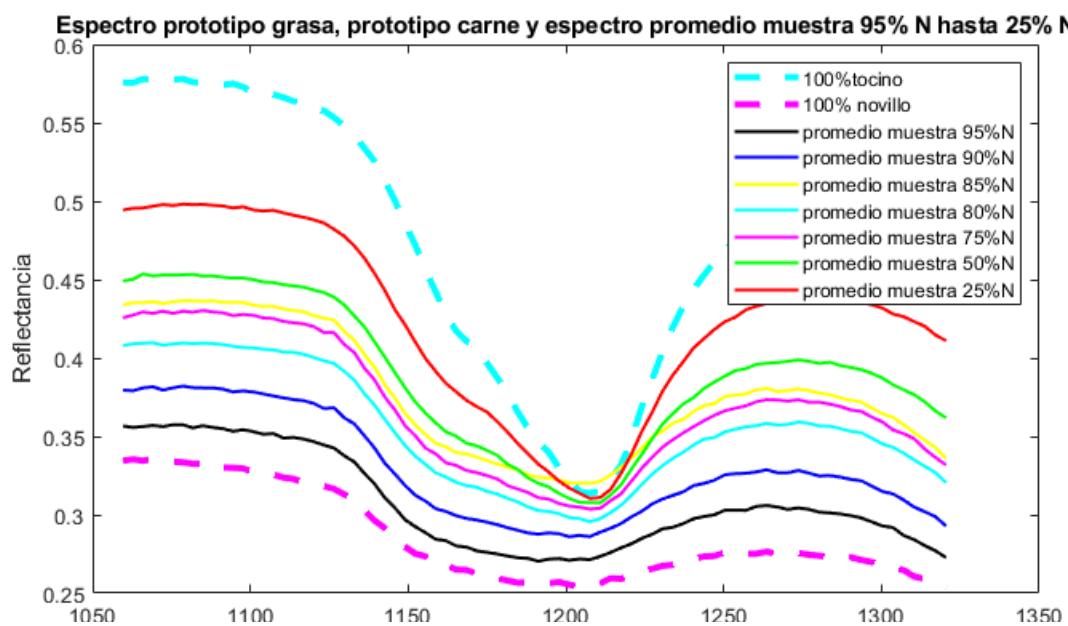


Figura 55: Espectros recortados en novilla SWIR para distancia euclídea

Como sucede en la región VIS-NIR, se vuelve a acotar el espectro (1050- 1350 nm) para obtener resultados de distancias euclídeas más precisos para facilitar la representación de los resultados de las distancias con la escala de colores. Como la zona de espectro que se acota comprende el pico de absorción de los lípidos a 1200nm la grasa de referencia sufre mucha absorción, siendo el punto donde más se junta con los promedios del resto de muestras. Como se puede apreciar en la Figura 56, las distancias no son muy grandes, por tanto, la distancia máxima toma el valor 2, y el mínimo muy próximo a 0. Observando la figura se ve que cuando una muestra cuando se compara con la referencia de carne, la representación gráfica se contrapone a la representación gráfica de esa muestra respecto a la grasa de referencia. Por ejemplo, en los puntos donde la distancia a la carne es mayor, luego en la gráfica de la grasa, se manifiesta con tonos asociados a distancias más pequeñas como cabría esperar.

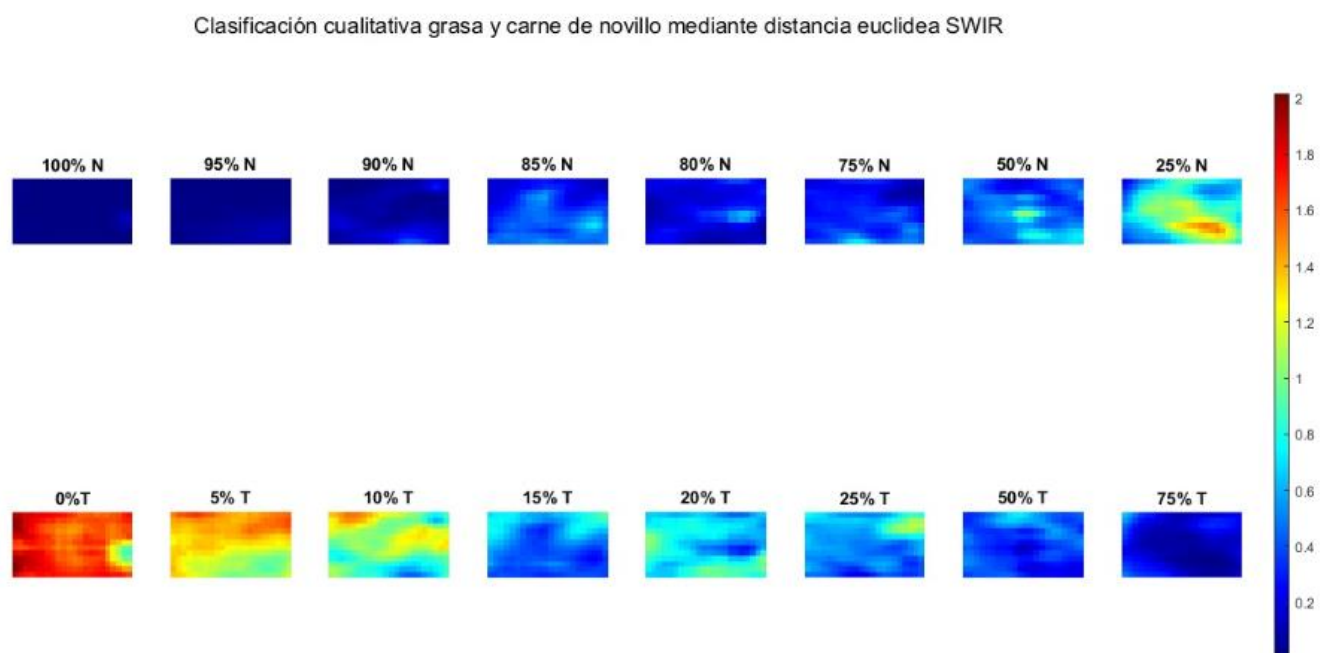


Figura 56: Clasificación cualitativa en novilla SWIR mediante distancia euclídea

### 7.2.2 Carne de cerdo

Se repite todo lo explicado anteriormente para la novilla, pero en este caso para la carne de cerdo. Para poder representar resultados concluyentes, se recorta el espectro y se obtienen las mismas formas que para el caso de la novilla, pero con valores de reflectancia más altos porque la carne de cerdo tiene valores de grasa intrínseca más altos.

### Región VIS-NIR

Los valores de la escala de referencia, entre el valor mínimo y máximo son semejantes a los de la carne de novilla en la región VIS-NIR. Observando la Figura 57 se ve cómo las distancias en la fila que representa la similitud con la carne, aumentan a medida que las mezclas contienen mayor porcentaje de grasa, y la distribución de los componentes en las distintas mezclas. En el caso de la grasa, se aprecia que a medida que aumenta el porcentaje de tocino, la distancia disminuye y se clasifica como grasa. En la gráfica se aprecian las variaciones en forma de degradado, desde las muestras con menos grasa a más grasa.

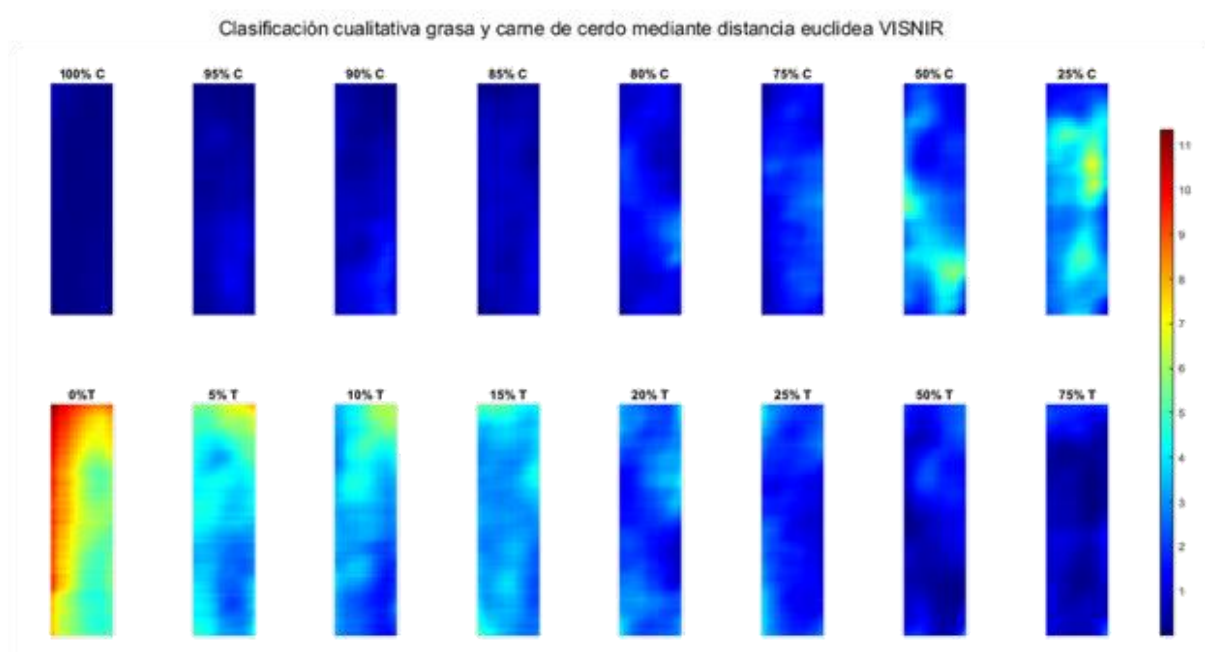


Figura 57: Clasificación cualitativa en cerdo VIS-NIR mediante distancia euclídea

### Región SWIR

Como en el caso de la novilla, por la zona del espectro que se recorta, donde la variación de los espectros promedios de las muestras es significativa, pero las distancias son más pequeñas, por lo tanto, los valores de la escala van desde el 0 hasta el 1.8 aproximadamente, que sería el valor de distancia máxima que se ha obtenido al recorrer todas las muestras y calcular todas las distancias.

Clasificación cualitativa grasa y carne de cerdo mediante distancia euclídea SWIR

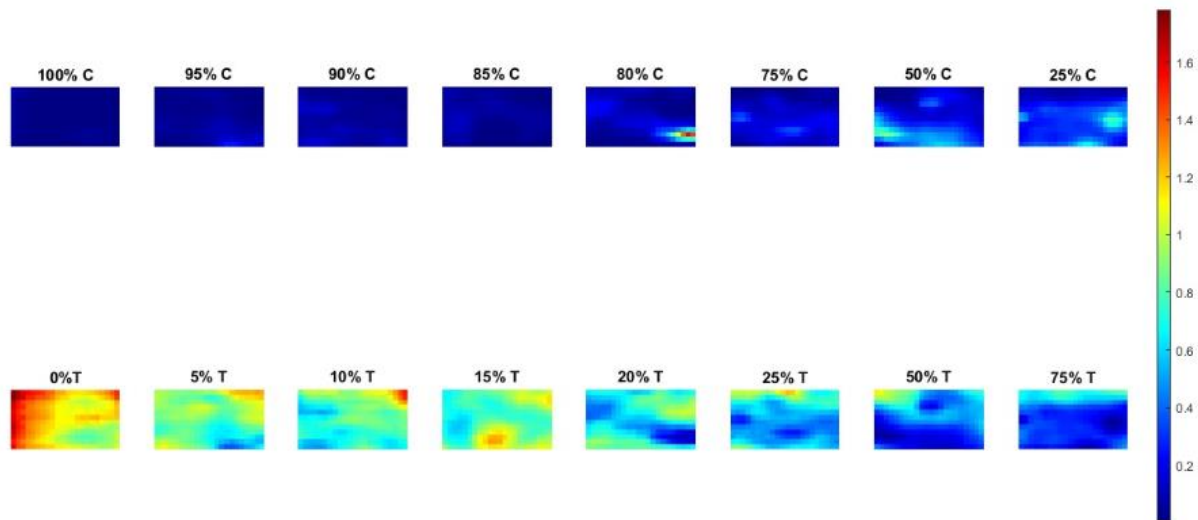


Figura 58: Clasificación cualitativa en cerdo SWIR mediante distancia euclídea

### 7.3 Clasificación mediante Mapeo del Ángulo Espectral (*SAM*)

Los principios de aplicación de este método son exactamente los mismos que para la distancia euclídea. En este caso, la comparación con los prototipos de referencia se lleva a cabo a través del ángulo que forman los vectores que se comparan. Para cada punto de las muestras se obtiene un vector con todos los valores del espectro, y se compara, por ejemplo, con la referencia de grasa a través del ángulo que forman. Cuanto más pequeño sea el ángulo más se parecen y, por tanto, se puede clasificar como grasa. Cuanto más grande, corresponderá a regiones de carne. El ángulo se consigue despejándolo de la fórmula del producto escalar entre los dos vectores. Los espectros en ambas regiones son exactamente iguales que las que se han representado en el proceso de la distancia euclídea, por lo que directamente se van a representar los resultados de aplicar la comparación del ángulo para clasificar las muestras.

#### 7.3.1 Carne de novilla

##### *Región VIS-NIR*

Se puede apreciar en este caso que los ángulos en cada punto de las muestras son mucho menores al compararlos con la referencia de carne. Por eso, la representación de la grasa, aunque existe una disminución del valor de los ángulos, y por tanto siendo más similar al vector de referencia, a medida que aumenta el porcentaje de grasa en las mezclas analizadas, los valores siguen siendo mucho mayores respecto al ángulo que forman el vector y la referencia de carne. La medida de muestra del 95% de novilla y 5% de tocino no cumplen el comportamiento general, y en ambos

casos cuando el ángulo es grande para la carne, también lo es para la grasa. Además, los valores tan altos de los ángulos en esa muestra comparando con la grasa hacen que el valor máximo de la escala de colores aumente dificultando observar el contraste entre las regiones en el resto de las muestras. Aunque las escalas para la representación gráfica sean diferentes porque los valores de los ángulos son mucho más pequeños que en la distancia euclídea, existe cierta correlación entre ambos métodos.

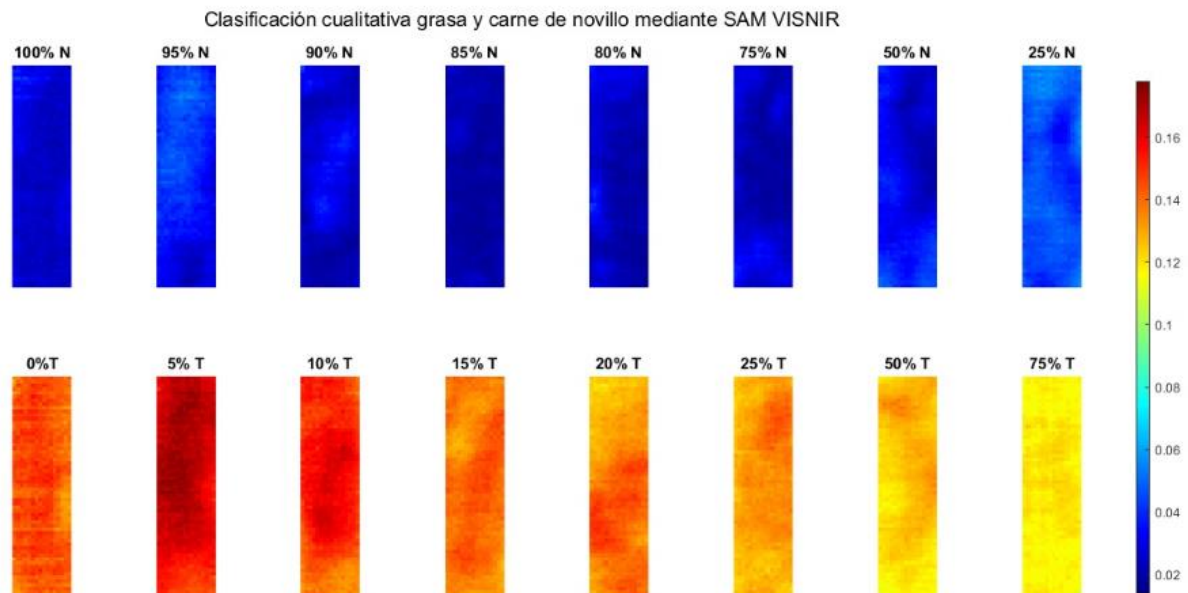


Figura 59: Clasificación cualitativa en novilla VIS-NIR mediante SAM

### Región SWIR

En este caso, la degradación es más evidente que en la región VIS-NIR. Esto quiere decir, por ejemplo, mirando la fila de la grasa, en las muestras con menor porcentaje de grasa los ángulos son grandes porque los espectros son muy dispares a la referencia de grasa, pero en la muestra de 75% de tocino se observa como el valor de los ángulos disminuye llegando a ser incluso menor que los ángulos comparando esa misma muestra con la referencia de carne (25% N).

Clasificación cualitativa grasa y carne de novillo mediante SAM SWIR

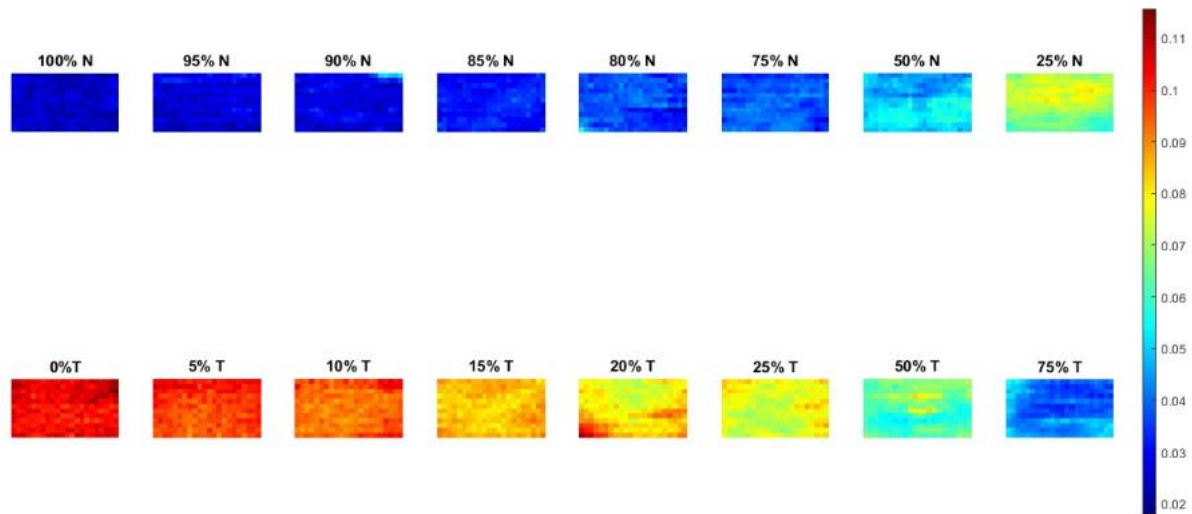


Figura 60: Clasificación cualitativa en novilla SWIR mediante SAM

### 7.3.2 Carne de cerdo

#### *Región VIS-NIR*

En este caso se puede observar mejor la variación de los porcentajes de grasa y carne en las muestras. Comparando los resultados obtenidos con los que se consiguen aplicando el método de la clasificación por distancia euclídea a estas mismas muestras se aprecia cierto nivel de correlación.

Clasificación cualitativa grasa y carne de cerdo mediante SAM VISNIR

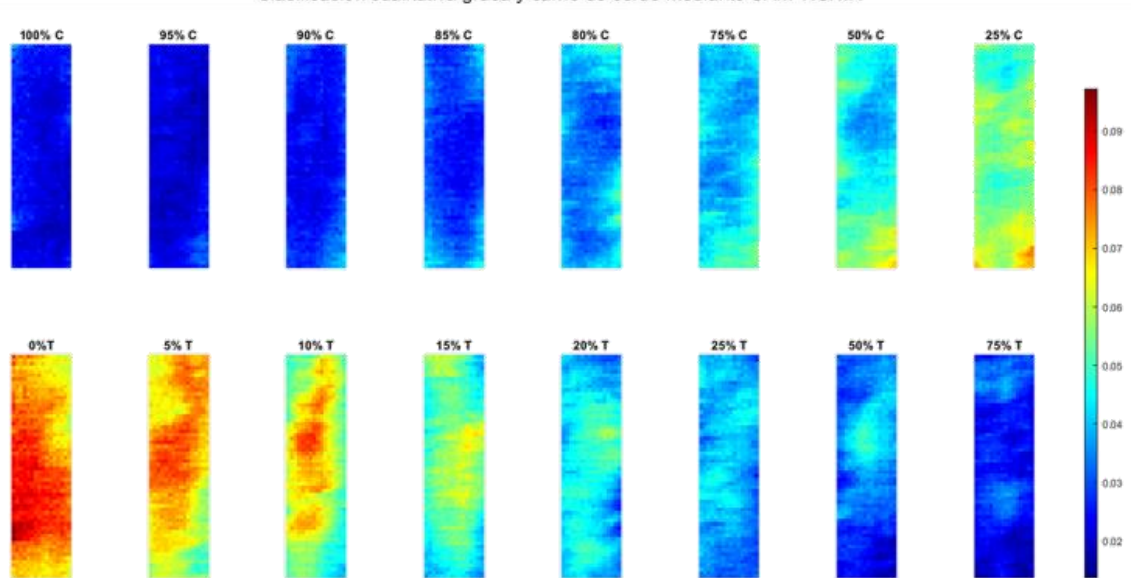
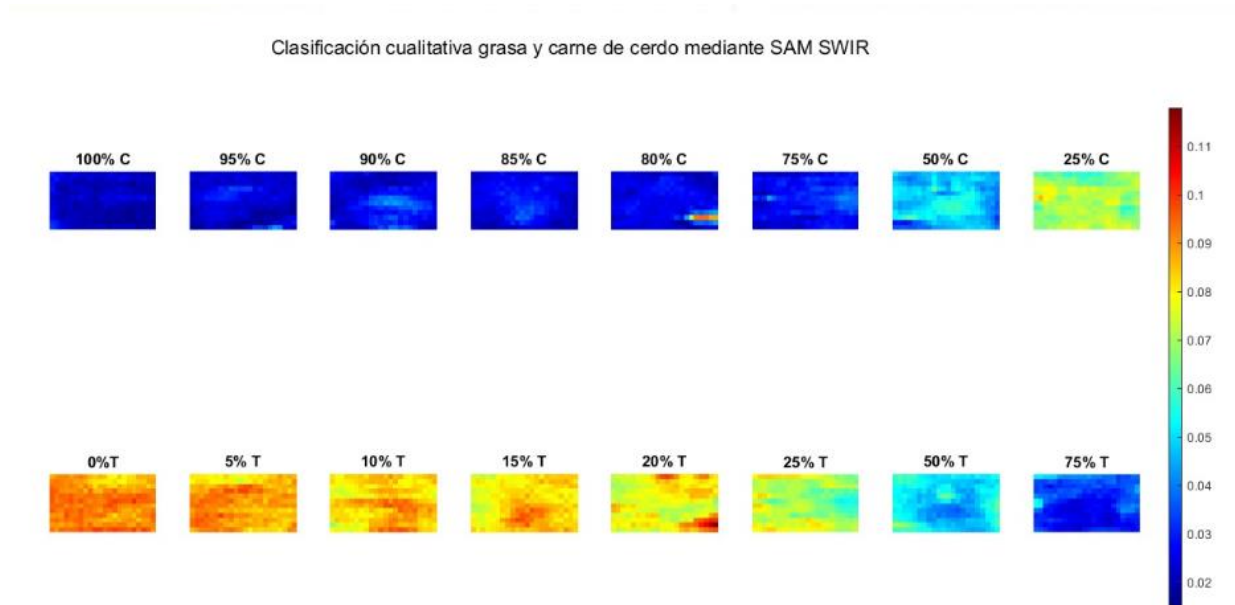


Figura 61: Clasificación cualitativa en cerdo VIS-NIR mediante SAM



### *Región SWIR*



**Figura 62:** Clasificación cualitativa en cerdo SWIR mediante SAM

Como sucede como norma general en todas las muestras, la variación de los ángulos comparados con la referencia de carne es menor que en el caso de la grasa, esto se puede deber que los espectros de las muestras correlan mejor con el espectro de referencia de carne 100% que con la forma del espectro de 100% de tocino.

Al realizar la comparación con la clasificación que realiza el método de la distancia euclídea para las mismas muestras y con el mismo sistema, también se aprecian similitudes en este caso, entre las regiones clasificadas como grasa o carne.



## Parte V: Conclusiones y líneas futuras





# Capítulo 8

## *Conclusiones generales y líneas futuras de investigación*

### 8.1 Conclusiones generales de este trabajo

A lo largo de este Trabajo de Fin de Grado se ha estudiado la aplicación de sistemas de imagen hiperespectral para el estudio de productos cárnicos, con el objetivo de poder caracterizar la composición química de las hamburguesas de una forma prácticamente instantánea, no invasiva y sin necesidad de destruir las muestras, de cara a mejorar los análisis de calidad y seguridad dentro del proceso de producción. Además, ofrece la posibilidad de proporcionar la información nutricional individualizada y precisa, determinando la proporción de los principales componentes que conforman la carne, como la grasa, y su distribución espacial dentro de la muestra.

El apartado teórico proporciona las bases para conocer los principios ópticos en los que se basan los sistemas espectroscópicos, el funcionamiento de estos y las principales características ópticas de los componentes que forman las hamburguesas. Muy importante esta última parte de cara a implementar las técnicas de procesado para obtener la información y poder interpretar correctamente los resultados.

La siguiente parte del trabajo incluye una descripción de las características técnicas de los sistemas utilizados (para regiones del espectro VIS-NIR y SWIR) y del proceso seguido a la hora de preparar las distintas mezclas de las muestras a medir, para una más fácil estimación caracterización de estas.

La metodología seguida se puede separar en dos bloques, dentro de las técnicas de procesado utilizadas, se han implementado dos métodos de clasificación cuantitativa, y otros tres en la parte de clasificación cualitativa. Todos ellos con el objetivo de cuantificar el porcentaje de carne y grasa de las hamburguesas.

Los métodos cuantitativos implementados son por un lado el método clásico de aproximación lineal por mínimos cuadrados (CLS, *Classical Least Squares*), método dependiente de la longitud de onda. El segundo método, PLS (*Partial Least Squares*) basado en procesos estadísticos de vectores propios. PLS permite estimar el porcentaje de grasa de cualquier hamburguesa después de haber entrenado el algoritmo con los espectros de las medidas con porcentajes conocidos, dando en la validación coeficientes de correlación altos, sobre todo en el sistema SWIR. En la región VIS-NIR, aunque con valores más bajos, la carne de cerdo proporciona una mejor respuesta. Analizando



ambos sistemas en conjunto, para cada tipo de carne se obtiene valores de correlación más bajo que realizando el análisis con cada sistema por separado.

El método clásico de mínimos cuadrados permite estimar el porcentaje de grasa a través de la aproximación lineal de los valores medidos para cada longitud de onda. Al calcular el grado de precisión de esa aproximación se obtienen las zonas del espectro más lineales, resultando no coincidir con los picos de absorción teóricos fundamentalmente en la región SWIR, donde en esas longitudes de onda para ambos tipos de carne, la correlación alcanza niveles muy bajos. En el estudio de este método se ha observado que el proceso de fabricación de las mezclas puede afectar en el cálculo del coeficiente de correlación del ajuste lineal si las muestras no contienen las proporciones adecuadas, desvirtuando el resultado obtenido al cuantificar una muestra con una composición desconocida como referencia, por lo tanto, el coeficiente de correlación mejora si se omite dicha muestra en el ajuste.

La parte de clasificación cualitativa se centra en conocer la distribución espacial de los componentes y comprobar así la homogeneidad de las muestras, pudiendo distinguir entre las regiones de carne y grasa. En el método que se basa en el algoritmo de umbralización Otsu, obtiene la certificación de la clasificación cualitativa de la grasa de las muestras a través de la correlación con la muestra de grasa de referencia medida. Se ha concluido que este método de clasificación sólo consigue resultados significativos midiendo con el sistema SWIR.

Otros métodos de clasificación cualitativa implementados son el cálculo de la distancia euclídea y la variación del ángulo entre vectores. En ellos se emplean como prototipos las muestras de referencia de la grasa y la carne a la hora de realizar las comparaciones y obtener las gráficas con la distribución espacial. El método de la distancia euclídea muestra mayor contraste en los resultados obtenidos, aunque ambos reflejan el degradado existente a medida que varían las proporciones en las muestras, al aumentar el porcentaje de grasa.

En todos los métodos se ha comparado la medida de la reflectancia medida directamente por el sistema con el valor de la reflectancia expresado mediante el logaritmo neperiano. Aunque se obtienen valores de correlación más bajos, expresar la reflectancia en logarítmico permite poder relacionar directamente este parámetro con la concentración de los principales cromóforos presentes en la carne.

Para finalizar, se puede concluir que todos los métodos implementados sirven para extraer información relevante de las medidas realizadas por ambos sistemas y caracterizar las muestras, o bien de forma cuantitativa o cualitativa, consiguiendo el objetivo de poder determinar los porcentajes de grasa y carne de las muestras de cerdo y novilla, además de conocer su distribución espacial.

## 8.2 Líneas futuras de trabajo

Este trabajo deja abiertas varias líneas de investigación, como por ejemplo aplicar técnicas de procesamiento complementarias para poder diferenciar mejor los picos de absorción más significativos y su magnitud mediante la implementación de la segunda derivada sobre los espectros promedio de las muestras, para poder obtener información más precisa de la concentración de los distintos cromóforos. Sería interesante conseguir mejor información en las zonas de interés donde el



coeficiente de correlación es muy bajo, como en los picos de absorción a 1200 nm y 1400 nm. Mediante la segunda derivada se podría diferenciar mejor que componentes intervienen en la absorción de esos puntos. Para aplicar la segunda derivada en la región SWIR, como se puede observar, primero hay que aplicar una técnica de suavizado de los espectros para limitar el ruido espectral y que el cálculo de las derivadas proporcione resultados concluyentes.

Otra posible línea de trabajo es comprobar el efecto del esparcimiento/*scattering* sobre las muestras y verificar si técnicas como el MSC (*Multiple Scatter Correction*) pudieran ayudar a mejorar los resultados.

Hay una gran variedad de técnicas de procesado que se podrían implementar para mejorar y complementar los resultados obtenidos, dependiendo de los aspectos que se quisieran investigar, por lo que la realización de este Trabajo de Fin de Grado sólo es el comienzo de un desarrollo más amplio de la aplicación de sistemas de imagen hiperespectral en productos cárnicos.



# Bibliografía

- [1] D.C. Ertugrul, “FoodWiki: a mobile app examines side effects of food additives via semantic web”, *Journal of Medical Systems*, (40-41) 2016.
- [2] Liu, Yi-Yun et al., “Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study”, *The Lancet Infectious Diseases*, 16(2), pp.161-168, 2016.
- [3] A. Nakonieczna et al., “Electrical impedance measurements for detecting artificial chemical additives in liquid food products”, *Food Control*, 66, pp.116-129, 2016.
- [4] D.W. Sun, “Infrared Spectroscopy for Food Quality Analysis and Control”, Elsevier, 2009.
- [5] D.Yang, Y.Ying, “Applications of Raman Spectroscopy in Agricultural Products and Food Analysis: A Review”, *Applied Spectroscopy Reviews*, 46, pp.539-560, 2011.
- [6] Y.S.Li, J.S.Church, “Raman spectroscopy in the analysis of food and pharmaceutical nanomaterials”, *Journal of Food and Drug Analysis*, 22, pp.29-48, 2014.
- [7] J. Sadecka, J.Tothova, “Fluorescence Spectroscopy and Chemometrics in the Food Classification, A Review, *Czech J. Food Sci.*, 25, pp.159-173, 2007.
- [8] A.Nawrocka, J.Lamorska, “Determination of Food Quality by Using Spectroscopic Methods”. *Advances in Agrophysical Research*, InTech Ed., 2013.
- [9] S.A.Baker, N.J.Miller-Ihli, P.Fodor, A.Woller, “Atomic Spectroscopy in Food Analysis”, ‘*Encyclopedia of Analytical Chemistry*’. John Wiley & Sons, Ltd, 2006.
- [10] P.S.Belton, R.Wood, “NMR Spectroscopy in Food Analysis”, *The Royal Society of Chemistry* (2012).



- [11] S.H.Baek, H.B.Lim, H.S.Chun, "Detection of Melamine in Foods Using Terahertz Time-Domain Spectroscopy", *Journal of Agricultural and Food Chemistry*, 62, pp.5403-5407, 2014.
- [12] W.Guo et al., "Effects of milk concentration and freshness on microwave dielectric properties", *Journal of Food Engineering*, 99, pp.344-350, 2010.
- [13] S.J.Lehotay, J.Hajslova, "Application of gas chromatography in food analysis", *Trends in Analytical Chemistry*, 21, pp.686-697, 2002.
- [14] P.M.Kus, S. van Ruth, "Discrimination of Polish unioral honeys using overall PTR-MS and HPLC fingerprints combined with chemometrics", *LWT - Food Science and Technology*, 62, pp.69-75, 2015.
- [15] H. Huang, H. Yu, Huirong Xu and Y. Ying, "Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: A review," *Journal of Food Engineering*, 87(3), pp.303-313, 2008.
- [16] H. Cen, Y. He, "Theory and application of near infrared reflectance spectroscopy in determination of food quality," *Trends in Food Science & Technology* 18(2), pp.72-83, 2007.
- [17] N.Prieto et al. "Application of near infrared reflectance spectroscopy to predict meat and meat product quality: A review", *Meat Science*, 83(2), pp.175-186, 2009.
- [18] OZAKI, Yukihiro; MCCLURE, W. Fred; CHRISTY, Alfred A. (ed.). "Near-infrared spectroscopy in food science and technology". John Wiley & Sons, 2006.
- [19] T. Isaksson, C. E. Miller, T. Næs. "Nondestructive NIR and NIT determination of protein, fat and water in plastic-wrapped, homogenized meat." *ApplSpectrosc*46:1685–1694, 1992.
- [20] OCU-Compra Maestra nº 378, feb.2013.
- [21] Ding, H. B., and R. J. Xu. "Near-infrared spectroscopic technique for detection of beef hamburger adulteration." *Journal of Agricultural and Food Chemistry* 48.6, 2193-2198, 2000.
- [22] Jacques, Steven L. "Optical properties of biological tissues: a review." *Physics in Medicine & Biology* 58.11, 2013.
- [23] Faber, Dirk J., et al. "Light absorption of (oxy-) hemoglobin assessed by spectroscopic optical coherence tomography." *Optics letters* 28.16, pp. 1436-1438, 2003.



- [24] Wilson, Robert H., et al. "Review of short-wave infrared spectroscopy and imaging methods for biological tissue characterization." *Journal of biomedical optics* 20.3, 030901, 2015.
- [25] Sekar, Sanathana Konugolu Venkata, et al. "Diffuse optical characterization of collagen absorption from 500 to 1700 nm." *Journal of biomedical optics* 22.1, 015006, 2017.
- [26] El Masry, Gamal, and Shigeki Nakauchi. "Prediction of meat spectral patterns based on optical properties and concentrations of the major constituents." *Food science & nutrition* 4.2, pp.269-283, 2016.
- [27] Jacquemoud, S., et al. "Estimating leaf biochemistry using the PROSPECT leaf optical properties model." *Remote sensing of environment* 56.3, pp.194-202, 1996.
- [28] R. L. P. van Veen, H. J. C. M. Sterenborg, A. Pifferi, A. Torricelli, and R. Cubeddu, "Determination of VIS- NIR absorption coefficients of mammalian fat, with time- and spatially resolved diffuse reflectance and transmission spectroscopy," in *Biomedical Topical Meeting*, OSA Technical Digest, Optical Society of America, paper SF4, 2004
- [29] Mehta, N.S., Shaik, S., Devireddy, R.V., & Gartia, M.R., "Single-Cell Analysis Using Hyperspectral Imaging Modalities", *Journal of biomechanical engineering*, 140 2, 2018.
- [30] S. Grusche, "Basic slit spectroscope reveals three-dimensional scenes through diagonal slices of hyperspectral cubes", 2014.
- [31] Gutiérrez-Gutiérrez, J. A., Pardo, A., Real, E., López-Higuera, J. M., & Conde, O. M. (2019). "Custom scanning hyperspectral imaging system for biomedical applications: modeling, benchmarking, and specifications", *Sensors*, 19(7), 1692.
- [32] Wang, *et al*, "Multiplexed Optical Imaging of Tumor-Directed Nanoparticles: A Review of Imaging Systems and Approaches", *Nanotheranostics*. 1. pp.369-388. 10.7150/ntno.21136, 2017.
- [33] J. A. Gutiérrez. "Integración RESTful de instrumentación de espectroscopia y su procesado como servicios", Trabajo de fin de máster, Universidad de Cantabria, 2018.
- [34] J.A Gutierrez, A. Pardo, J.M López Higuera, O.M Conde, "Application of rotary mirror hyperspectral imaging systems for agrifood and biomedical scenarios", 2019



- [35] Lu, *et al*, "Innovative Hyperspectral Imaging-Based Techniques for Quality Evaluation of Fruits and Vegetables: A Review", *Applied Sciences*. 7. 189. 10.3390/app7020189, 2017.
- [36] J. Workman, "Applied Spectroscopy: A Compact Reference for Practitioners", *Optical Spectrometers*, p. 3-28, 1997.
- [37] Eguizabal, Alma, *et al*, "Direct identification of breast cancer pathologies using blind separation of label-free localized reflectance measurements." *Biomedical optics express*, 4.7, pp. 1104-1118, 2013.